What to Learn from a Few Visible Transitions' Statistics?

Pedro E. Harunari,[1,2,*] Annwesha Dutta,[3,4] Matteo Polettini,[2] and Édgar Roldán[3,†]

[1]*Instituto de Física da Universidade de São Paulo, 05314-970 São Paulo, Brazil*
[2]*Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg, Luxembourg*
[3]*ICTP—The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34151 Trieste, Italy*
[4]*Department of Physics, Indian Institute of Science Education and Research, Tirupati 517507, India*

(Received 26 March 2022; revised 17 August 2022; accepted 29 September 2022; published 7 December 2022)

Interpreting partial information collected from systems subject to noise is a key problem across scientific disciplines. Theoretical frameworks often focus on the dynamics of variables that result from coarse-graining the internal states of a physical system. However, most experimental apparatuses can only detect a partial set of transitions, while internal states of the physical system are blurred or inaccessible. Here, we consider an observer who records a time series of occurrences of one or several transitions performed by a system, under the assumption that its underlying dynamics is Markovian. We pose the question of how one can use the transitions' information to make inferences of dynamical, thermodynamical, and biochemical properties. First, elaborating on first-passage time techniques, we derive analytical expressions for the probabilities of consecutive transitions and for the time elapsed between them, which we call intertransition times. Second, we derive a lower bound for the entropy production rate that equals the sum of two non-negative contributions, one due to the statistics of transitions and a second due to the statistics of intertransition times. We also show that when only one current is measured, our estimate still detects irreversibility even in the absence of net currents in the transition time series. Third, we verify our results with numerical simulations using unbiased estimates of entropy production, which we make available as an open-source toolbox. We illustrate the developed framework in experimentally validated biophysical models of kinesin and dynein molecular motors, and in a minimal model for template-directed polymerization. Our numerical results reveal that while entropy production is entailed in the statistics of two successive transitions of the same type (i.e., repeated transitions), the statistics of two different successive transitions (i.e., alternated transitions) can probe the existence of an underlying disorder in the motion of a molecular motor. Taken all together, our results highlight the power of inference from transition statistics ranging from thermodynamic quantities to network-topology properties of Markov processes.

DOI: 10.1103/PhysRevX.12.041026 Subject Areas: Statistical Physics

I. INTRODUCTION

Model systems in physics [1], chemistry [2–4], biology [5–7], and computation [8] are routinely described by Markov processes, which are also amenable to thermodynamic analysis [9–13]. This approach thrives when there is full knowledge of the system's internal state, but in most practical applications, experimental apparatuses access few degrees of freedom or have a finite resolution; thus, only partial information is available. One example is the rotation of flagella in a bacterial motor [14]: Observation of orientation switches in the direction of the bacteria's flagella suggests the existence of internal states that are hidden from the observer.

The problem of measuring partial information, or of coarse-graining degrees of freedom, is usually framed in terms of the internal state of a system [15–20]. However, in most practical applications, an external observer only measures "footprints" of one or several transitions, rather than the internal state itself, as sketched in Fig. 1(a). These footprints may be due to physical degrees of freedom satisfying microscopic reversibility, in which case it is possible to talk about their energetic and entropic balance, as sketched in Fig. 1(b), where the observer can detect the emission and absorption of a photon γ, or the production or consumption of a chemical species X. Finally, Fig. 1(c) sketches the motion of a molecular motor (e.g., a kinesin) along a periodic track (e.g., microtubule). The motor

*pedroharunari@gmail.com
†edgar@ictp.it

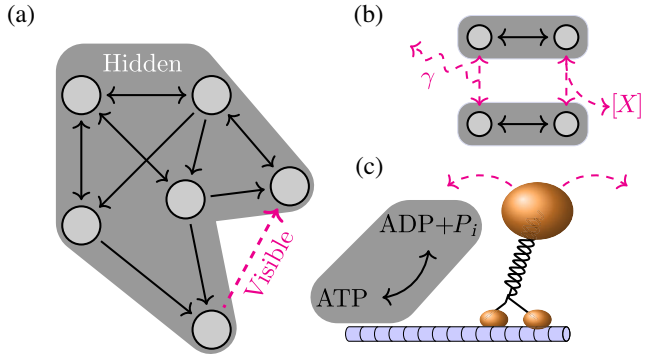


FIG. 1. Illustrations of the partial information acquired by an apparatus that can only detect a few visible transitions. (a) Setup of our framework, for the case of only one visible transition. A physical system performs a Markov-jump process in a network of states (circles), all of which are hidden from an external observer. The observer can only see occurrences of a few visible transitions (one, in this example), while the rest of the transitions remain hidden during the data acquisition (black shaded area). (b,c) Model examples described by our theory: (b) photon emission or absorption γ and synthesis or consumption of chemical species X that signal the occurrence of some transitions, and (c) a molecular motor performing steps along a track and transitions related to spatial motion along the track, detected by monitoring the position of the cargo (orange sphere), with chemical fuel consumption (ATP hydrolysis) often remaining hidden (gray box); see also Fig. 2.

undergoes structural changes followed by a translocation step associated with the consumption of some resources (e.g., adenosine triphosphate [ATP]). In this case, the only visible transitions are the forward and backward steps along the track. As explained below, this situation is customary in experiments where the motion of a microscopic bead attached to the motor can be used to detect spatial displacements along the track while conformational changes and chemical fuel consumption remain undetectable to the experimenter [21].

Significant developments in single-molecule experimental techniques with biological systems at cellular and subcellular levels have been reported over the last few decades [22]. For example, the motion of biomolecular machines involved in cellular transport—such as kinesin [23], dynein [24,25], and myosin [26]—has been resolved at the subnanometer resolution. Examples include real-time tracking of individual, fluorescently tagged biomolecules [27,28] followed by data analysis techniques of the recorded trajectories using, e.g., kymographs [29,30]. In most of these experiments, biomolecular machines are subject to nonequilibrium forces that may be intrinsic (e.g., chemical reactions) or extrinsic (e.g., mechanical forces exerted by optical tweezers). This motivates the fact that the motion of the molecular motor is routinely described by Markovian nonequilibrium stationary states.

The typical scenario of single-molecule studies is such that only a partial set of degrees of freedom and/or transitions are experimentally accessible. For example, using high-resolution optical tweezers, it is customary that the spatial transitions (e.g., a step in a linear track) can be measured experimentally while conformational changes or chemical reactions remain hidden from the experimenter. This is the case of, e.g., the molecular machines of the central dogma of genetic information processing, DNA polymerase [22], RNA polymerase [31], and ribosomes [32,33]. Because every transition during molecular motor motion is accompanied by changes in internal energy due to the chemical energy arising from the coupling of the system to chemical reservoirs [34,35], having reliable estimates of entropy production from the observation of a partial set of transitions is key to developing accurate bounds on efficiency and thermodynamic costs of molecular machines [36,37].

In an attempt to extract useful thermodynamic information from the partial observation of a few visible transitions' statistics, we develop a transition-based coarse-graining framework for continuous-time Markov processes. Our analytical progress leads to descriptions and predictions suitable for systems whose available information comes only from counting transitions and measuring the time elapsed between two consecutive transitions—a key concept that we denote as intertransition times. In particular, we focus on how one can infer thermodynamic and topological properties from the sole observation of intertransition times and frequencies of transitions, and what the consequences are for experimentally validated models of biomolecular systems.

II. GOALS AND MAIN RESULTS

Recent work revealed that information extracted from transitions between a few selected visible states provides information about entropy production [38–40]. Yet, most of these efforts relied on knowledge about the internal states of the system. Instead, the main question we address in this contribution is the following: What can be learned about a system solely from the occurrence of a few visible transitions (denoted $\ell_i \in \mathcal{L}$) and from the time elapsed between them (denoted t_i and called intertransition time)? Our object of study is therefore a time series of the form

$$\Gamma_\tau^{\mathcal{L}}: \quad \vdash\!\!\overset{t_0}{\text{---}}\ \ell_0\ \overset{t_1}{\text{---}}\ \ell_1\ \overset{t_2}{\text{---}}\ \cdots\ \overset{t_{n+1}}{\text{---}}\!\!\dashv \qquad (1)$$

where t_i denotes the time elapsed between the occurrence of two successive transitions $\ell_{i-1}, \ell_i \in \mathcal{L}$, with ℓ_0 the first transition observed. Notice that the subindex $\tau = \sum_{i=0}^{n+1} t_i$ in $\Gamma_\tau^{\mathcal{L}}$ indicates the total time duration of the observed trajectory, which is a deterministic quantity, whereas t_i are all positive random variables.

In the following, we use Dirac's notation for vectors, where $|i\rangle$ is a column vector with entries $\delta_{i,j}$ for j spanning through the state space; thus, for example, $\langle i|\mathbf{A}|j\rangle = A_{ij}$ is

the ith row and jth column entry of matrix $\mathbf{A}$. We introduce a special notation when we deal with transitions $\ell \in \mathcal{L}$: Here, $\langle\langle \ell \|$ is a row vector that has all zero entries except the element corresponding to the source state of transition ℓ. On the other hand, $\|\ell\rangle\rangle$ is a column vector that has all zero entries except the element corresponding to the target state of transition ℓ. For example, transition $\ell = 1 \to 3$ has $\langle\langle \ell \| = \langle 1|$ and $\|\ell\rangle\rangle = |3\rangle$, and the matrix element associated with transition $1 \to 3$ is $\langle\langle \ell \|\mathbf{A}^{\mathsf{T}}\|\ell\rangle\rangle = \langle 3|\mathbf{A}|1\rangle = A_{3,1}$, where T denotes matrix transposition [41].

We assume that the underlying (hidden) dynamics that produces the collected data is a continuous-time, discrete-state-space Markov process with time-independent rates (also known as a jump process) over an irreducible network (from now on, simply called the Markov chain). The time series $\Gamma_\tau^{\mathcal{L}}$ is reminiscent of so-called hidden Markov processes, but we emphasize again that here our focus is on visible transitions rather than visible states. We focus on the following statistical quantities, which are easily accessible in experimental settings:

(i) Histograms collecting the frequency $F(t|\ell_i, \ell_{i+1})dt$ that the time t_i elapsed between ℓ_i and ℓ_{i+1}, called the intertransition time, which lies within the interval $[t, t + dt)$.

(ii) The conditional frequency $F(\ell_{i+1}|\ell_i)$ that a transition ℓ_{i+1} is observed, given that the previous was ℓ_i. We call the case $\ell_i = \ell_{i+1}$ repeated transitions and the case $\ell_i \neq \ell_{i+1}$ alternated transitions.

(iii) The frequency that a transition ℓ occurs in an observed trajectory $F(\ell)$.

In this paper, we characterize these quantities from a statistical, a thermodynamic, and a biophysical point of view. The first task (statistical) is important from a fundamental point of view, to understand which features from a hidden process can be learned by looking only at the statistics of a few visible transitions. The second (thermodynamic) task is focused on inferring the rate of entropy production of the underlying Markov chain, which is a key quantity to characterize the irreversibility of a nonequilibrium process. The third (biophysical) task is important from an applied point of view because most single-molecule experiments retrieve partial information about the nonequilibrium dynamics of biological systems.

To tackle these objectives, we derive analytical expressions for the expected value of the three aforementioned transition statistics. From the thermodynamic point of view, on the additional assumption that for every visible transition $\ell \in \mathcal{L}$ its reversed $\bar{\ell} \in \mathcal{L}$ is also visible—which we dub visible reversibility—we compute and characterize the visible stationary rate of entropy production,

$$\sigma_{\mathcal{L}} = \lim_{\tau\to\infty} \frac{1}{\tau} D(P[\Gamma_\tau^{\mathcal{L}}]||P[\bar{\Gamma}_\tau^{\mathcal{L}}]), \tag{2}$$

defined as the rate of Kullback-Leibler divergence [42]

$$D(P[\Gamma_\tau^{\mathcal{L}}]||P[\bar{\Gamma}_\tau^{\mathcal{L}}]) = \int \mathcal{D}\Gamma_\tau^{\mathcal{L}} P[\Gamma_\tau^{\mathcal{L}}] \ln(P[\Gamma_\tau^{\mathcal{L}}]/P[\bar{\Gamma}_\tau^{\mathcal{L}}]) \tag{3}$$

of the probability density of $\Gamma_\tau^{\mathcal{L}}$ with respect to that of its suitably defined time-reversed trajectory $\bar{\Gamma}_\tau^{\mathcal{L}}$. Finally, we apply the formalism to stochastic models of the molecular motor motion of dynein, kinesin, and polymerization in disordered tracks.

Our main results are as follows:

(1) *Analytical expressions for intertransition-time probabilities in terms of parameters of the hidden Markov chain*. To this aim, we solve analytically a first-passage time problem in transition space, i.e., a "first-transition time" problem [47]. More specifically, letting $\mathbf{W}$ be any transition rate matrix (generator of a Markov chain), we introduce a survival matrix $\mathbf{S}$ obtained by setting the entries in $\mathbf{W}$ corresponding to the visible transitions to zero [see Eq. (14) for a rigorous definition]. Mapping the occurrence of transitions to a first-passage-time problem, for the probability density of transition ℓ_{i+1} happening in the infinitesimal time interval $[t, t + dt)$, and given that the previous visible transition was ℓ_i, we find

$$P(t, \ell_{i+1}|\ell_i) = -\langle\langle \ell_{i+1}\|\mathbf{W}^{\mathsf{T}}\|\ell_{i+1}\rangle\rangle \times \langle\langle \ell_{i+1}\| \exp(t\mathbf{S})\|\ell_i\rangle\rangle. \tag{4}$$

The first factor $\langle\langle \ell_{i+1}\|\mathbf{W}^{\mathsf{T}}\|\ell_{i+1}\rangle\rangle$ is the rate of transition ℓ_{i+1}, and the second factor $\langle\langle \ell_{i+1}\| \exp(t\mathbf{S})\|\ell_i\rangle\rangle$ is the probability of going from state $\|\ell_i\rangle\rangle$ to $\langle\langle \ell_{i+1}\|$ in time t without performing any visible transition.

From Eq. (4), we obtain an explicit expression for the conditional probability of successive transitions, the intertransition-time probability density, and the probability of the next observed transition given the current occupation distribution. Furthermore, we provide explicit expressions in the case of hidden state spaces with ring topology, which we validate with the above analytical expression (4). Equation (4) generalizes results in first-passage time problems from reaching a subset of states [48,49] to performing an arbitrary subset of transitions.

(2) *Assuming visible reversibility, the visibility of the opposite of each visible transition*. We calculate the stationary rate of entropy production $\sigma_{\mathcal{L}}$ given by Eq. (2) and compare it with that of the hidden Markov chain, σ. In particular, we prove that

$$\sigma_{\mathcal{L}} \leq \sigma, \tag{5}$$

with the equality holding for systems with ring topology or for systems in which every single transition is visible.

Furthermore, we also show that the visible entropy production rate can be written as the sum of two independent contributions,

$$\sigma_{\mathcal{L}} = \sigma_{\ell} + \sigma_t, \tag{6}$$

both of which are positive because they take the form of Kullback-Leibler divergences of transition statistics, i.e., $\sigma_{\ell} \geq 0$ and $\sigma_t \geq 0$. The contribution σ_{ℓ} depends solely on the mere occurrence of transitions, whereas σ_t depends on the observed intertransition times. Analytical expressions for σ_{ℓ} and σ_t can be found in Eqs. (52) and (53), given here for convenience:

$$\sigma_{\ell} = \langle K \rangle \sum_{\ell,\ell' \in \mathcal{L}} P(\ell|\ell') P(\ell') \ln \frac{P(\ell|\ell')}{P(\overline{\ell'}|\bar{\ell})}, \tag{7}$$

$$\sigma_t = \langle K \rangle \sum_{\ell,\ell' \in \mathcal{L}} P(\ell|\ell') P(\ell') \times D[P(t|\ell',\ell)||P(t|\bar{\ell},\overline{\ell'})], \tag{8}$$

where $\langle K \rangle$ is the visible traffic rate, i.e., the expected number of visible transitions that occur over time [50], sometimes also called dynamical activity [51]. The sums in Eqs. (7) and (8) run over the set of visible transitions $\mathcal{L}$, and the bar in $\bar{\ell}$ denotes the opposite direction of ℓ, i.e., the observed transition when the dynamics is time reversed. In Eq. (8) and in the following, we denote by $P(t|\ell',\ell)$ the probability density for the intertransition time between ℓ' followed by ℓ, and we also introduce a key quantity given by the Kullback-Leibler divergence between intertransition-time distributions

$$D[P(t|\ell',\ell)||P(t|\bar{\ell},\overline{\ell'})] = \int_0^{\infty} dt P(t|\ell',\ell) \ln \frac{P(t|\ell',\ell)}{P(t|\bar{\ell},\overline{\ell'})}. \tag{9}$$

For the relevant case of only two visible transitions in forward ("+") and backward ("−") directions between the same pair of states, i.e., $\mathcal{L} = \{+,-\}$, Eqs. (7) and (8) simplify to Eqs. (54) and (55), given here for convenience:

$$\sigma_{\ell} = \langle K \rangle [P(+) - P(-)] \ln \frac{P(+|+)}{P(-|-)}, \tag{10}$$

$$\sigma_t = \langle K \rangle P(+|+) P(+) D[P(t|+,+)||P(t|-,-)] + \langle K \rangle P(-|-) P(-) D[P(t|-,-)||P(t|+,+)]. \tag{11}$$

Interestingly, both depend on the statistics of *repeated* transitions; σ_{ℓ} depends on the conditional probabilities $P(\ell|\ell)$, and σ_t depends on intertransition-time probability densities $P(t|\ell,\ell)$ through the Kullback-Leibler divergences

$$D[P(t|+,+)||P(t|-,-)] = \int_0^{\infty} dt P(t|+,+) \ln \frac{P(t|+,+)}{P(t|-,-)},$$
$$D[P(t|-,-)||P(t|+,+)] = \int_0^{\infty} dt P(t|-,-) \ln \frac{P(t|-,-)}{P(t|+,+)}. \tag{12}$$

The value of $\sigma_{\mathcal{L}}$ allows us to improve on entropy production rate estimates previously proposed [39,52]. We also show that our approach provides a tighter bound for σ than some of the so-called thermodynamic uncertainty relations [53,54], especially in situations where the net current is small (e.g., for molecular motors close to stall force, the force at which the motor stops moving).

(3) *Application of the formalism to three distinct stochastic models in cell biology: motion of dynein and kinesin on linear tracks, and template-directed polymerization processes in the presence of disorder.* Particularly interesting from these examples is the finding that intertransition times of repeated $P(t|\ell,\ell)$ and alternate transitions [viz. $P(t|\ell,\bar{\ell})$] carry different information about the hidden Markov chain. Whereas repeated transitions allow us to estimate dissipation, alternated transitions provide hints about disorder.

The paper is structured as follows: In Sec. III, we develop our framework and derive Eq. (4) for generic Markov chains; in Sec. IV, we obtain the results for a pair of transitions in opposite directions along a system with ring topology; in Sec. V, we consider transitions over a pair of states to address the problem of estimation of entropy production; in Sec. VI, we discuss biophysical applications for dynein, kinesin, and motion in disordered tracks; finally, we conclude with a discussion in Sec. VII. Detailed mathematical proofs are given in the appendixes. Results similar to those in the present paper are discussed in the companion article [55]; see Sec. VII for a more detailed discussion.

III. VISIBLE TRANSITIONS' STATISTICS

A. Framework

We consider continuous-time Markov chains over a finite and discrete state space $\{1,2,\ldots,N\}$. We assume that the state-space structure is such that any two states are connected by only one transition and that the network of states is irreducible. Thus, we assume that there always

exists a nonzero probability path from any to every state. The Perron-Frobenius theorem ensures the existence of a unique stationary distribution, towards which the system relaxes, and the system's ergodicity, the equivalence between time and ensemble averages. The occupation probability at time t is expressed as a column vector $|p(t)\rangle = (p_1(t), p_2(t), \ldots, p_N(t))^{\mathsf{T}}$ obeying the master equation

$$\frac{d}{dt}|p(t)\rangle = \mathbf{W}|p(t)\rangle, \tag{13}$$

where $\mathbf{W}$ is a time-independent stochastic matrix with positive nondiagonal elements W_{ij}, which are the transition rates from state j to i, and negative diagonal elements $W_{ii} = -\sum_{j\neq i} W_{ij}$ are the escape rates from state i.

An observer unambiguously detects transitions that belong to a subset $\mathcal{L}$ of all possible transitions, while the remaining transitions and the occupancy of internal states go unnoticed. Visible transitions $\ell \in \mathcal{L}$ connect state $\langle\langle \ell||$ to a different state $||\ell\rangle\rangle$. In jump processes, transitions are instantaneous, and the system spends time in states, called sojourn times. We define the intertransition time as the sum of all sojourn times between two consecutive visible transitions.

We introduce the survival matrix $\mathbf{S}$, obtained by subtracting from the stochastic matrix the transition rates related to every visible transition:

$$\mathbf{S} \equiv \mathbf{W} - \sum_{\ell\in\mathcal{L}} ||\ell\rangle\rangle\langle\langle \ell||\mathbf{W}^{\mathsf{T}}||\ell\rangle\rangle\langle\langle \ell||, \tag{14}$$

where the term being summed is a matrix with all zero entries but for term $\langle\langle \ell||\mathbf{W}^{\mathsf{T}}||\ell\rangle\rangle$, which is the rate of transition ℓ.

B. Main results

The survival propagator $\exp(t\mathbf{S})$ describes the system's evolution given that no visible transition occurs. It does not conserve probability because not every column of $\mathbf{S}$ adds up to zero; thus, it can be interpreted as a transition matrix of a process with probability leakages whenever a transition in $\mathcal{L}$ takes place.

Consider a succession of transitions and intertransition times, as in Eq. (1), and create the histogram of times conditioned on the occurrence of the previous and next transitions. This provides the empirical definition of the intertransition time's frequency:

$$F_\tau(t|\ell_i, \ell_{i+1})dt \sim \text{histogram}(t|\ell_i, \ell_{i+1}). \tag{15}$$

The frequency that the next observed transition is ℓ_{i+1} given that the previous is ℓ_i can be obtained as

$$F_\tau(\ell_{i+1}|\ell_i) \equiv \frac{\#(\ell_i \to \ell_{i+1})}{\sum_{j=1}^{|\mathcal{L}|} \#(\ell_i \to \ell_j)}, \tag{16}$$

where $\#(\ell_i \to \ell_{i+1})$ is the number of transitions ℓ_i followed by ℓ_{i+1} and $|\mathcal{L}|$ is the number of visible transitions, the cardinality of subset $\mathcal{L}$. Furthermore, the frequency that one observed transition is ℓ_i among all transitions in a trajectory is

$$F_\tau(\ell_i) \coloneqq \frac{\#\ell_i}{\sum_{j=1}^{|\mathcal{L}|} \#\ell_j}. \tag{17}$$

Because of the system's ergodicity, all empirical probabilities have as both expected and asymptotic values the real probability, $P(\cdot) = \langle F_\tau(\cdot)\rangle = \lim_{\tau\to\infty} F_\tau(\cdot)$.

Finding the probability that by time t the system has not performed any transitions in $\mathcal{L}$ and then performs ℓ is a first-transition time problem whose solution leads to our main result below.

Result: Let $\mathbf{W}$ be the transition matrix of a continuous-time and stationary discrete-state-space irreducible Markov chain and consider a subset $\mathcal{L}$ of all possible transitions and the survival matrix as in Eq. (14). The joint probability that the intertransition time falls within $[t, t+dt)$ and that the next visible transition is $\ell_{i+1} \in \mathcal{L}$, given that the last observed transition was $\ell_i \in \mathcal{L}$, is

$$P(t, \ell_{i+1}|\ell_i)dt = \langle\langle \ell_{i+1}||\mathbf{W}^{\mathsf{T}}||\ell_{i+1}\rangle\rangle\langle\langle \ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle dt, \tag{18}$$

in agreement with Ref. [55]. See Appendix A for a proof.

All other probabilities we are interested in can be obtained from Eq. (18), whose joint probability can be split into $P(t, \ell_{i+1}|\ell_i) = P(\ell_{i+1}|\ell_i)P(t|\ell_i, \ell_{i+1})$. The conditional probability of the next observed transition can be obtained by integrating over time, resulting in

$$\begin{aligned} P(\ell_{i+1}|\ell_i) &= \int_0^\infty dt P(\ell_{i+1}, t|\ell_i) \\ &= -\langle\langle \ell_{i+1}||\mathbf{W}^{\mathsf{T}}||\ell_{i+1}\rangle\rangle\langle\langle \ell_{i+1}||\mathbf{S}^{-1}||\ell_i\rangle\rangle. \end{aligned} \tag{19}$$

Without the need for additional assumptions, the probability density of intertransition time t between such a transition and its preceding one can be obtained by dividing the joint probability by the transition probability above,

$$P(t|\ell_i, \ell_{i+1}) = \frac{P(t, \ell_{i+1}|\ell_i)}{P(\ell_{i+1}|\ell_i)} = -\frac{\langle\langle \ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle}{\langle\langle \ell_{i+1}||\mathbf{S}^{-1}||\ell_i\rangle\rangle}. \tag{20}$$

It satisfies $\int_0^\infty dt P(t|\ell_i, \ell_{i+1}) = 1$ for all ℓ_i and ℓ_{i+1}.

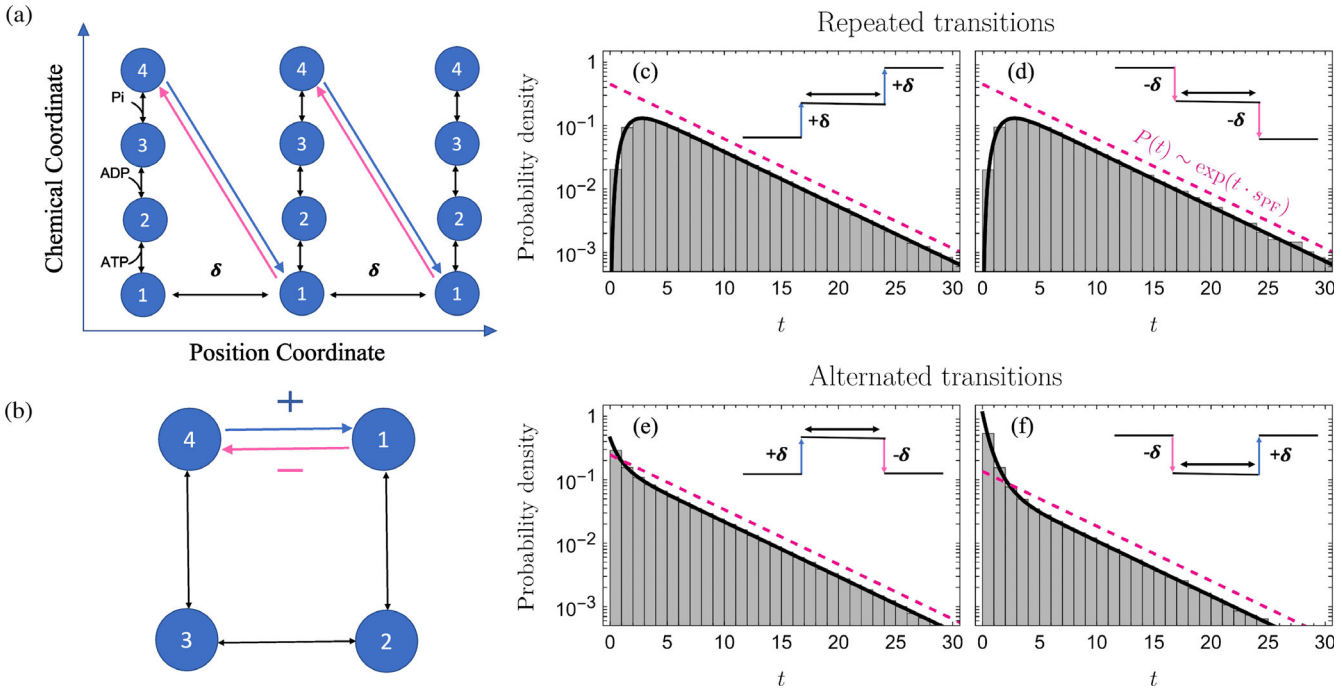


FIG. 2. Illustration of transition statistics in an ATP-driven motion of a molecular machine in which only spatial displacements are visible to an experimental apparatus. Chemical transitions (ATP binding $1 \to 2$, ADP release $2 \to 3$, and Pi release $3 \to 4$) are hidden from the apparatus. (a,b) Illustration of the model state network, given by a four-state model (b) arranged along a periodic spatial lattice (a) along which the motor moves only when executing the $4 \to 1$ transition (motion forward $+$ with step size δ) or the $1 \to 4$ transition (motion backward $-$ with step size δ). (c)–(f) Histograms for the intertransition-time probability densities obtained from Gillespie simulations (bars), and analytical predictions given by Eq. (20). The dashed magenta lines have slope given by the largest eigenvalue of the survival operator $\mathbf{S}$ [see Eq. (25)], and the inserted diagrams are illustrations of the displacement of the motor during the respective pair of transitions (see insets).

We name the expected number of visible transitions over time $\langle K \rangle$ the visible traffic rate, inspired by time-symmetric quantities relevant in the analysis of stochastic systems far from equilibrium [56], sometimes referred to as dynamical activity [51] and frenesy [50,57,58]. Analytically, its value can be obtained in the limit $t \to \infty$ as

$$\langle K \rangle = \sum_{\ell \in \mathcal{L}} \langle\langle \ell || \mathbf{W}^{\mathsf{T}} || \ell \rangle\rangle \langle\langle \ell || p_\infty \rangle\rangle, \tag{21}$$

where $|p_\infty\rangle$ is the stationary distribution given by the solution of $\mathbf{W}|p_\infty\rangle = 0$. Furthermore, the stationary probability that a visible transition is $\ell \in \mathcal{L}$ is given by

$$P(\ell) = \frac{1}{\langle K \rangle} \langle\langle \ell || \mathbf{W}^{\mathsf{T}} || \ell \rangle\rangle \langle\langle \ell || p_\infty \rangle\rangle. \tag{22}$$

C. Numerical illustration of the framework

Figure 2 presents an example of the application of our approach to a model of a molecular motor with four internal states that are driven by the consumption of ATP [59]. The motor performs spatial displacements along a filament through the only visible transition in a single-molecule experiment. Figure 2(a) is a scheme of the motor's motion; transitions in the chemical coordinate involve consumption and production of chemical species and are considered invisible for the experimenter; conversely, transitions in the position coordinate $1 \leftrightarrow 4$ are considered visible since they result in spatial displacement of size δ (mechanical movement), and in this case, they compose the subset $\mathcal{L}$. Figure 2(b) shows the irreducible network in which a Markov chain describes the evolution, and visible transitions $4 \to 1 \equiv +$ and $1 \to 4 \equiv -$ are, respectively, related to forward and backwards displacement. Figures 2(c)–2(f) show an excellent agreement between numerical simulations and Eq. (20) for all the distributions of intertransition times between repeated $(++,--)$ and alternated $(+-, -+)$ transitions. While alternated transitions yield an intertransition-time distribution that is monotonously decreasing, the distribution of intertransition times between repeated transitions is nonmonotonous. This is because of network topology constraints: Whereas for alternated transitions $t = 0$ is the most likely event, repeated transitions require motion over the entire hidden network, which renders the probability of $t = 0$ almost impossible for large hidden networks. Furthermore, we observe that the distributions of all intertransition times have the same exponential tail [magenta dashed line in Figs. 2(c)–2(f). This is consistent with theory, as discussed in the next subsection.

D. Additional remarks

(i) *Moments of the intertransition times.* Equation (20) is key for further results of this work. An immediate outcome is that, since Eq. (20) is the probability density of the intertransition time of ℓ_{i+1} after ℓ_i, the mean intertransition time can also be obtained from the survival matrix:

$$\int_0^\infty \mathrm{d}t\, tP(t|\ell_i,\ell_{i+1}) = -\frac{\langle\langle \ell_{i+1}||(\mathbf{S}^{-1})^2||\ell_i\rangle\rangle}{\langle\langle \ell_{i+1}||\mathbf{S}^{-1}||\ell_i\rangle\rangle}, \quad (23)$$

and higher-order moments can be obtained analogously.

(ii) *Generalization of first-passage times.* First-passage times between states can be obtained as a particular case of first-transition times. Let $\mathcal{L}_i = \{i' \to i\colon \forall\ i' \neq i\}$ be the set of all transitions leading to an absorbing state i. The first time that state i is reached coincides with the first time that one of these transitions is observed. From Eq. (18), we then obtain the probability density for the first-passage time of reaching state i starting from a state $j \neq i$,

$$\begin{aligned} P(t,i|j)dt &= \sum_{\ell\in\mathcal{L}_i} P(t,\ell|j)dt \\ &= \sum_{\ell\in\mathcal{L}_i} \langle\langle \ell||\mathbf{W}^{\mathsf{T}}||\ell\rangle\rangle\langle\langle \ell||\exp(t\mathbf{S})|j\rangle dt, \end{aligned} \quad (24)$$

where $P(t,\ell|j)dt$ is the probability that, starting from j, the first visible transition observed in the time interval $[t, t+dt)$ is ℓ. This latter result is well known; see, e.g., Ref. [1].

(iii) *Connection to large deviation theory.* Consider the number of times $\#_t(\ell)$ a transition is performed up to time t (sometimes called flux, or counting field). Notice that $\#_t(\ell)$ only vanishes for all $\ell \in \mathcal{L}$ if no visible transition has been performed. Therefore, the generating function of its moments $\langle e^{\lambda\sum_{\ell\in\mathcal{L}}\#_t(\ell)}\rangle$ in the limit $\lambda \to -\infty$ is precisely the survival probability density. The moment-generating function can be calculated as $\sum_y \langle y|\exp(t\mathbf{W}_\lambda)|x\rangle$ [60], where $\mathbf{W}_\lambda$ is the so-called tilted matrix, which in the limit $\lambda \to -\infty$ reduces to $\mathbf{S}$, consistently with Eq. (18).

(iv) *Existence of* $\mathbf{S}^{-1}$. Since the process defined by $\mathbf{W}$ is ergodic, for a large enough time, the system will perform at least one of the observed transitions with probability 1: $\lim_{t\to\infty}|p(t)\rangle = \lim_{t\to\infty}\exp(t\mathbf{S})|p(0)\rangle = \vec{0}$, where $\vec{0}$ is a vector of zeros. This is ensured by the fact that every eigenvalue of $\mathbf{S}$ has a negative real part, $\mathrm{Re}(\lambda_i) < 0\ \forall\ i$, as proved in the Supplementary Material of Ref. [61]. Such a property also guarantees the convergence of the integral $\int_0^\infty \mathrm{d}t\exp(t\mathbf{S})$, which is required to normalize the probability in Eq. (20), and $\det(\mathbf{S}) = \prod_i \lambda_i \neq 0$, which grants the existence of $\mathbf{S}^{-1}$.

(v) *Probability of instantaneous pairs.* The propagator acting over a state results in a probability vector with non-negative entries, $\exp(t\mathbf{S})||\ell_i\rangle\rangle \geq 0$; therefore, $\partial_t\langle\langle \ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle = \langle\langle \ell_{i+1}||\mathbf{S}\exp(t\mathbf{S})||\ell_i\rangle\rangle$ has the same sign as $\langle\langle \ell_{i+1}||\mathbf{S}||\ell_i\rangle\rangle$ at $t = 0$. If the transition ℓ_{i+1} starts in the same state where ℓ_i ended, $\langle\langle \ell_{i+1}|| = ||\ell_i\rangle\rangle$, the intertransition time has nonvanishing probability of being zero since the diagonal entries of $\partial_t\langle\langle \ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle$ are always negative. Conversely, for sequences of transitions with $\langle\langle \ell_{i+1}|| \neq ||\ell_i\rangle\rangle$, the null intertransition time has zero probability: The observer has to wait for internal jumps to occur before ℓ_{i+1} takes place. This property explains the shape of intertransition-time probability densities in Fig. 2: For alternated transitions $+-$ and $-+$, the source state of the second transition is the target of the first transition; therefore, the probability of instantaneous intertransition time is nonzero [cf. panels (d) and (e)]. On the other hand, for repeated transitions $++$ and $--$, instantaneous intertransition times cannot be realized because one needs to perform additional transitions [cf. panels (c) and (f)].

(vi) *Universality of the tails.* Notice that it is always possible to decompose the numerator in Eq. (20) as $\langle\langle \ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle = \sum_{k=1}^N c_k e^{ts_k}$, where s_k are the eigenvalues of $\mathbf{S}$ and c_k are real coefficients obtained by projecting onto its eigenvectors, under the assumption that $\mathbf{S}$ has a nondegenerate spectrum, and with minor modifications of the argument otherwise [62]. Assuming hidden irreducibility, i.e., the irreducibility of the state space after the removal of all visible transitions, this property implies that the long-time behavior of the intertransition-time distribution is independent of the visible transitions ℓ_{i+1} and ℓ_i:

$$\lim_{t\to\infty}\frac{1}{t}\ln\left[\langle\langle \ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle\right] \asymp s_{\mathrm{PF}}, \quad (25)$$

where s_{PF} is the dominant Perron-Frobenius root, a negative real value. Therefore, all intertransition-time distributions have the same exponential tail given by the largest eigenvalue of $\mathbf{S}$. This can be observed in Figs. 2(c)–2(f), where tails of the histograms obtained for the four types of intertransition times match the value given by s_{PF}.

IV. EXPLICIT RESULTS FOR UNICYCLIC NETWORKS

In addition to the developed generic framework, analytical expressions for intertransition-time distributions in

ring (unicyclic) networks can be obtained using Laplace transforms. To this aim, we now use a combinatoric graph-theoretic approach based on sums over all possible hidden paths in Laplace space. As shown below, these explicit calculations showcase that computing intertransition statistics in generic Markov chains is often a Herculean task, which is greatly simplified by the exact analytical framework developed in Sec. III.

In a variety of models of, e.g., enzymatic reactions [63], the state space can be depicted as a ring network, where every state i is connected to only its nearest neighbors $i \pm 1$ and nothing else ($1 \leftrightarrow 2 \leftrightarrow \ldots \leftrightarrow N \leftrightarrow 1$). In particular, we consider as visible the pair of transitions between states $1 \leftrightarrow N$, without loss of generality. For this section, we also assume that every neighboring state has transitions in both directions, $W_{i,i+1} > 0$ and $W_{i+1,i} > 0$, allowing for the cycle performance in both orientations.

We denote the two visible transitions as follows: The clockwise transition from state N to 1 is $+ \equiv N \to 1$, and the counterclockwise transition is $- \equiv 1 \to N$. There are four possible intertransition times to be considered, between pairs of successive transitions $++$, $+-$, $-+$, and $--$.

The probability density of spending time τ_j in a given state j before the next transition to i (often called sojourn time) is given by

$$\pi_{ij}(\tau_j) \coloneqq P(\tau_j, j \to i|j) = W_{ij} \exp(-W_{jj}\tau_j). \quad (26)$$

To characterize the different paths that intertwine the desired transitions, we introduce the number of times the pair of opposite transitions between $i \leftrightarrow i+1$ are performed as k_i, and the number of possible paths satisfying $\vec{k} = (k_1, \ldots, k_{N-1}, k_N = 0)$ is $C^{\vec{k}}_{+,+}$. The simplest "bare" path leading to $++$ is a sequence of clockwise transitions starting and ending in 1; the probability density of performing it in an interval t, up to a normalization constant, is given by the convolution of sojourn times,

$$P_{\text{bare}}(t|+,+) \propto \int d\vec{\tau} \pi_{1,N}(\tau_N) \cdots \pi_{2,1}(\tau_1) \delta\left(t - \sum_i \tau_i\right), \quad (27)$$

where δ is the Dirac delta distribution. As the Laplace transform of convolutions is the product of Laplace transforms, we further deal with products of terms in the form

$$\hat{\pi}_{ij}(s) \equiv \int_0^\infty d\tau_j \pi_{ij}(\tau_j) e^{-s\tau_j} = \frac{W_{ij}}{W_{jj} + s}, \quad (28)$$

where the hat $\hat{\cdot}$ denotes the Laplace transform and, for simplicity, we often suppress the dependency on s, the complex frequency corresponding to time in the Laplace space.

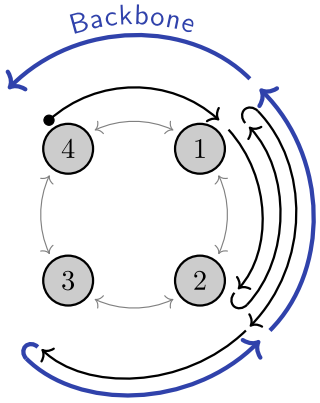


FIG. 3. Illustration of a single trajectory in a four-state Markov process where only transitions $4 \to 1$ (+) and $1 \to 4$ (−) are visible. Thin gray arrows represent the possible transitions in state space, while thick arrows represent the trajectory. For this example, we define its associated "backbone" (blue thick arrows) by the sequence of the transitions last traveled between each pair of states that were visited in the trajectory: $3 \to 2$, $2 \to 1$, and $1 \to 4$. See text for further details.

To count all trajectories, we solve a nontrivial combinatoric problem introducing the concept of backbone: For a given path, its associated backbone is composed of the set of every last transition performed between each pair of visited states, see Fig. 3. Once the backbone associated with a path is identified, all other variables in the trajectory can freely change without changing the fact that the trajectory starts and ends at two prescribed visible transitions. For example, in the case of repeated transitions $++$, a path characterized by $\vec{k}$ will contain k_1 transitions $2 \to 1$ and $k_1 + 1$ transitions $1 \to 2$. The last transition performed, $1 \to 2$, ensures that the path is moving in the direction of eventually performing $+$ again and is part of the backbone; the rest of the backbone will come from transitions $2 \to 3$, $3 \to 4$, and so on.

A. Repeated transitions

For the case $++$, notice that each pair of states $i \leftrightarrow i+1$ accommodates $2k_i$ transitions in a path, half clockwise and half counterclockwise, and then one extra transition that belongs to the backbone, ensuring that the path is not stuck between these two states. The intertransition-time probability density can be obtained from the convolution of every sojourn time in a path and by summing over all possible paths. Its Laplace transform is given by

$$\hat{P}(s|+,+) = \frac{1}{\mathcal{N}_{+,+}} \overbrace{\left(\prod_{i=1}^{N} \hat{\pi}_{i+1,i}\right)}^{\text{backbone}} \times \sum_{k_1,\ldots,k_{N-1}=0}^{\infty} C^{\vec{k}}_{+,+} \prod_{i=1}^{N-1} \underbrace{[\hat{\pi}_{i,i+1}\hat{\pi}_{i+1,i}]^{k_i}}_{\equiv x_i}, \quad (29)$$

where we define $x_i = \hat{\pi}_{i,i+1}\hat{\pi}_{i+1,i}$ as the product of Laplace transformed sojourn times in two opposite directions and impose the condition $N+1 \equiv 1$.

The initial-value theorem for Laplace transforms states that $P(0) = \lim_{s\to\infty} s\hat{P}(s)$; therefore, the constant $\mathcal{N}_{+,+}$ can be obtained by $\hat{P}(0|+,+) = 1$, which ensures that the inverse Laplace transform $P(t|+,+)$ is normalized. The combinatorial coefficient $C^{\vec{k}}_{+,+}$ is the number of all possible paths between $++$ with transitions satisfying $\vec{k}$. It is obtained in Appendix B and reads

$$C^{\vec{k}}_{+,+} = \prod_{i=2}^{N-1} \binom{k_i + k_{i-1}}{k_i}. \tag{30}$$

Equation (29) simplifies by plugging in Eq. (30) and introducing a continued fractions generator $\Theta[x_i] \coloneqq x_i/(1-\Theta[x_{i-1}])$ that truncates at $\Theta[x_1] = x_1$. From the property $\sum_{k=0}^{\infty}\binom{n+k}{k}x^k = (1-x)^{-(n+1)}$, valid for $|x| < 1$, we obtain a simplified expression,

$$\hat{P}(s|+,+) = \frac{1}{\mathcal{N}_{+,+}}\hat{\pi}_{1,N}\prod_{i=1}^{N-1}\frac{\hat{\pi}_{i+1,i}}{1-\Theta[x_i]}. \tag{31}$$

The case $--$ can be obtained analogously upon the substitutions $i \mapsto N-i+1$, $\forall\, i \in [1,N]$, and $\Xi[x_i] \coloneqq x_i/(1-\Xi[x_{i+1}])$, with $\Xi[x_{N-1}] = x_{N-1}$:

$$\hat{P}(s|-,-) = \frac{1}{\mathcal{N}_{-,-}}\hat{\pi}_{N,1}\prod_{i=1}^{N-1}\frac{\hat{\pi}_{i,i+1}}{1-\Xi[x_i]}. \tag{32}$$

By a diagrammatic approach to explicitly obtain the continued fraction generators Θ and Ξ (Appendix C), we find that intertransition-time densities are the same for repeated transitions,

$$P(t|+,+) = P(t|-,-). \tag{33}$$

Such a property is reminiscent of the so-called generalized Haldane equality [63–65], which states that the probability density of waiting time t until a system performs a clockwise cycle, given that a counterclockwise cycle was not performed, is the same as waiting time t for the opposite phenomenon. This property can be observed in Figs. 2(d) and 2(e) and, in general, is not satisfied for alternated transitions.

Since $\lim_{s\to\infty}\hat{\pi}_{ij} = 0$ for every pair i, j, applying the initial-value theorem to Eqs. (31) and (32) results in

$$P(0|+,+) = P(0|-,-) = 0, \tag{34}$$

apart from very specific choices of transition rates that might forbid the existence of the limit. This result confirms that instantaneously performing a full cycle has zero probability and gives a characteristic shape to the histograms in Figs. 2(c) and 2(f).

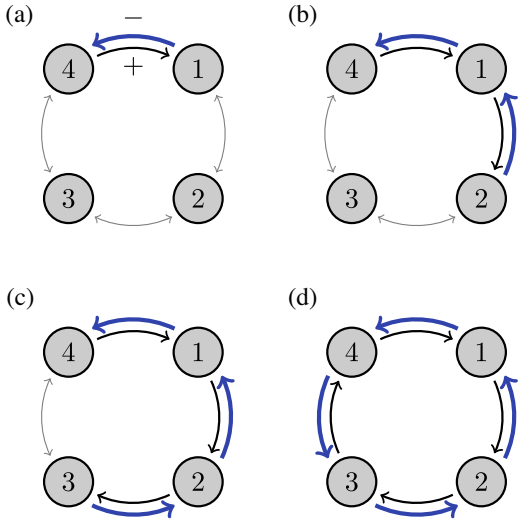


FIG. 4. Illustration of the four possible backbones (thick blue arrows) that ensure the completion of a $(-)$ $1 \to 4$ transition after a $(+)$ $4 \to 1$ transition, for the example of the four-state Markov model shown in Fig. 3. Thick gray arrows represent the possible transitions, while thick arrows represent the transitions used by trajectories of different lengths. From panels (a)–(d), the trajectories have $M = 0$ to $M = 3$. More details are given in the main text.

B. Alternated transitions

Between alternated transitions, it is not necessary to cover the whole state space. In fact, it is possible to not have any transitions in between the visible ones, and this is how a zero intertransition time might occur; thus, there is no analogue of Eq. (34) for alternated transitions.

In this case, there are N possible backbones; they are composed of $M+1 \in [1,N]$ transitions with the same orientation starting from the farthest visited state to the target of the last visible transition; see Fig. 4. Also, we observe an even number $2k_i$ of transitions between pairs $i \leftrightarrow i+1$.

The Laplace transform of the intertransition-time probability density for the pair $+-$ is

$$\hat{P}(s|+,-) = \frac{1}{\mathcal{N}_{+,-}}\hat{\pi}_{N,1}\sum_{M=0}^{N-1}\sum_{k_1,\ldots,k_M=1}^{\infty} C^{\vec{k},M}_{+,-}\prod_{i=1}^{M}[\underbrace{\hat{\pi}_{i,i+1}\hat{\pi}_{i+1,i}}_{\equiv x_i}]^{k_n}, \tag{35}$$

where the backbone contributions come from the sum over M and $\hat{\pi}_{N,1}\prod_{i=1}^{M}\hat{\pi}_{i,i+1}$.

The coefficient

$$C^{\vec{k},M}_{+,-} = \prod_{i=2}^{M}\binom{k_i + k_{i-1} - 1}{k_i} \tag{36}$$

counts the number of possible paths leading to $+-$ with a given $\vec{k}$ and backbone length of M (more details can be found in Appendix B). Once again, from the property $\sum_{k=0}^{\infty}\binom{n+k}{k}x^k=(1-x)^{-n-1}$, we obtain a simplified expression

$$\hat{P}(s|+,-)\propto\hat{\pi}_{N,1}\sum_{M=0}^{N-1}\frac{\prod_{i=1}^{M}x_i}{\prod_{j=1}^{M-1}(1-\Xi[x_j])^2}\frac{1}{1-\Theta[x_M]}, \quad (37)$$

and analogously, we find

$$\hat{P}(s|-+)\propto\hat{\pi}_{1,N}\sum_{M=0}^{N-1}\frac{\prod_{i=N-M}^{N-1}x_i}{\prod_{j=N-M+1}^{N-1}(1-\Xi(x_j))^2}\times\frac{1}{1-\Xi(x_{N-M})}. \quad (38)$$

Applying the initial-value theorem to Eqs. (37) and (38) results in

$$P(0|+,-)=\frac{W_{N,1}}{\mathcal{N}_{+,-}} \quad (39)$$

and

$$P(0|-,+)=\frac{W_{1,N}}{\mathcal{N}_{-,+}}. \quad (40)$$

We recall that $\mathcal{N}_{+,-}$ and $\mathcal{N}_{+,-}$ can be obtained by $\hat{P}(0|\bullet)=1$ as a property of Laplace transforms since $P(t|\bullet)$ is normalized. The nonvanishing contribution comes from the terms with $M=0$, which means that it is possible to instantaneously observe a pair of alternated transitions, and it is due to the shortest backbone of all: a single transition. This can be observed in the shape of histograms in Figs. 2(d) and 2(e).

To obtain the intertransition-time densities, one needs to perform an inverse Laplace transform on Eqs. (31), (32), (37), and (38). We remark that, while possible, in general, it is not straightforward to find closed analytical expressions to such inverse Laplace transforms (cf. Appendix B of Ref. [66]). Notice that it is possible to obtain all moments of intertransition times without resorting to the inverse Laplace transforms by using the relation

$$\langle t^n|\ell_i,\ell_{i+1}\rangle=(-1)^n\left[\frac{\partial^n}{\partial s^n}\hat{P}(s|\ell_i,\ell_{i+1})\right]_{s=0}. \quad (41)$$

V. IRREVERSIBILITY AND ENTROPY PRODUCTION

Entropy production and time irreversibility are the thermodynamic footprints of nonequilibrium dynamics. In stochastic thermodynamics, irreversibility of nonequilibrium stationary processes can be quantified by the asymmetry between a process and its time reversed in terms of the Kullback-Leibler divergence of forward to backward probabilities [44] that provide bounds for the rate of entropy production. As we now show, in a jump process, this asymmetry is present in the sequence of visited states and, also, in the intertransition times. In this section, we introduce an inference scheme for the entropy production rate of a system for which only a few transitions are visible. We also assume visible reversibility, i.e., that every visible transition can be performed in its opposite direction, and the opposite of a visible transition is also visible. This scenario is typical in physical settings such as electron hopping between leads or a molecular motor walking along a microtubule.

The stationary rate of entropy production in the system plus environment is a measure of time-reversal asymmetry in the dynamics of the system averaged over all microscopic trajectories γ_τ over state space:

$$\sigma=\lim_{\tau\to\infty}\frac{1}{\tau}\underbrace{\sum_{\gamma_\tau}P[\gamma_\tau]\ln\frac{P[\gamma_\tau]}{P[\bar{\gamma}_\tau]}}_{D(P[\gamma_\tau]||P[\bar{\gamma}_\tau])}, \quad (42)$$

where $\bar{\gamma}_\tau$ is the time-reversed trajectory obtained by reverting in time the states visited along the trajectory γ_τ. For Markovian nonequilibrium time-independent processes, it has been shown [67] that the entropy production (42) depends only on the statistics of jumps between different states as follows:

$$\sigma=\sum_{i<j}J_{ij}\ln\frac{W_{ij}}{W_{ji}}, \quad (43)$$

where

$$J_{ij}=W_{ij}p_j(\infty)-W_{ji}p_i(\infty) \quad (44)$$

denotes the stationary probability current from state j to state i [68] and, for convenience, we have set the Boltzmann constant k_B to unity. Currents can be empirically observed when the involved transitions are visible. In other words, if $\ell=j\to i$, the current can be empirically obtained by

$$J_{ij}=\lim_{\tau\to\infty}\frac{\#\ell-\#\bar{\ell}}{\tau}, \quad (45)$$

where we recall that τ is the trajectory duration and # represents the number of occurrences.

The entropy production rate from the available data in the present framework is obtained by comparing visible trajectories $\Gamma_\tau^{\mathcal{L}}$ that can be seen as a transition-based coarse-graining of the full trajectory γ_τ over state space. A key result for our estimates is the chain rule for the Kullback-Leibler divergence between two random variables [43],

which has also been applied to stochastic processes [44,69]: $D[\rho_1(x,y)||\rho_2(x,y)] \geq D[\rho_1(x)||\rho_2(x)]$ for any two distributions ρ_1 and ρ_2 of two random variables x and y. Because the trajectories $\Gamma_\tau^{\mathcal{L}}$ contain less random variables than the microscopic trajectories γ_τ (e.g., most of the transitions and their associated intertransition times are not included in $\Gamma_\tau^{\mathcal{L}}$), one gets $D(P[\gamma_\tau]||P[\bar{\gamma}_\tau]) \geq D(P[\Gamma_\tau^{\mathcal{L}}]||P[\bar{\Gamma}_\tau^{\mathcal{L}}])$, which implies the inequality

$$\sigma_{\mathcal{L}} = \lim_{\tau\to\infty} \frac{1}{\tau} D(P[\Gamma_\tau^{\mathcal{L}}]||P[\bar{\Gamma}_\tau^{\mathcal{L}}]) \leq \sigma. \tag{46}$$

In the following, we employ the transition information using $\sigma_{\mathcal{L}}$ to obtain lower bounds for the entropy production, and we analyze how $\sigma_{\mathcal{L}}$ can be computed in practice from simulations or experimental data.

A. Inference of entropy production

We ask the question of how the inferred entropy production rate $\sigma_{\mathcal{L}}$ can be computed in practice and study how tight the lower bound is. The observer collects a coarse-grained trajectory during an interval $[0,\tau]$ comprising visible transitions in both directions $\ell \in \mathcal{L}$ and the intertransition times between them:

$$\Gamma_\tau^{\mathcal{L}} = \{(\ell_0, t_0), (\ell_1, t_1), \ldots, (\ell_n, t_n)\}, \tag{47}$$

with $\sum_{i=0}^{n+1} t_i = \tau$. The construction of the time-reversed trajectory in the transition space requires special care. The time-reversed trajectory is given by the sequence of reversed transitions $\bar{\ell}$ in the opposite order, and the intertransition times are shifted: If the time before a transition ℓ_i is t_i, in the time-reversed dynamics, the time before transition $\bar{\ell}_i$ is t_{i+1}; see Fig. 5 for an illustrative example. Thus,

$$\bar{\Gamma}_\tau = ((\bar{\ell}_n, t_{n+1}), (\bar{\ell}_{n-1}, t_n), \ldots, (\bar{\ell}_0, t_1)). \tag{48}$$

The probability of a trajectory can be written in terms of the conditional probabilities of consecutive transitions and waiting times $P(t_i, \ell_i|\ell_{i-1}) = P(t_i|\ell_{i-1}, \ell_i)P(\ell_i|\ell_{i-1})$; hence,

$$P[\Gamma_\tau^{\mathcal{L}}] = P(t_0, \ell_0)P(t_1, \ell_1|\ell_0)\cdots P(t_n, \ell_n|\ell_{n-1}), \tag{49}$$

$$P[\bar{\Gamma}_\tau] = P(t_{n+1}, \bar{\ell}_n)P(t_n, \bar{\ell}_{n-1}|\bar{\ell}_n)\cdots P(t_1, \bar{\ell}_0|\bar{\ell}_1). \tag{50}$$

After working out these expressions (see Appendix D for details), we derive the following decomposition of the irreversibility measure $\sigma_{\mathcal{L}}$:

$$\sigma_{\mathcal{L}} = \sigma_\ell + \sigma_t, \tag{51}$$

where the first term is the contribution from the sequence of transitions,

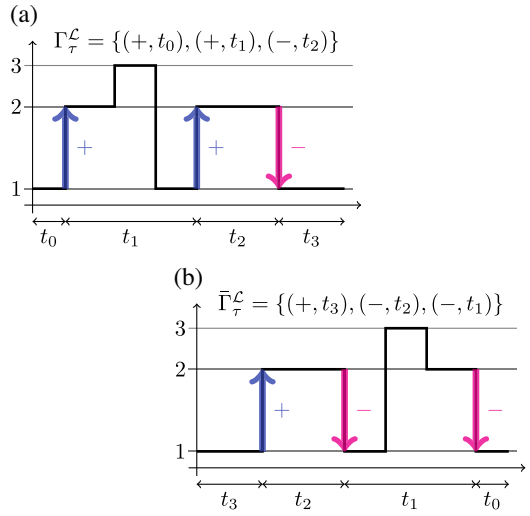


FIG. 5. (a) Hidden trajectory over states $\{1,2,3\}$ and the observation $\Gamma_\tau^{\mathcal{L}}$ for $\tau = t_0 + t_1 + t_2 + t_3$ and $\mathcal{L} = \{+:1\to 2, -:2\to 1\}$. (b) Time-reversed hidden trajectory and resulting $\bar{\Gamma}_\tau^{\mathcal{L}}$.

$$\sigma_\ell = \langle K\rangle \sum_{\ell,\ell'\in\mathcal{L}} P(\ell|\ell')P(\ell') \ln \frac{P(\ell|\ell')}{P(\bar{\ell}'|\bar{\ell})}, \tag{52}$$

and the second from the intertransition times

$$\sigma_t = \langle K\rangle \sum_{\ell,\ell'\in\mathcal{L}} P(\ell|\ell')P(\ell')D[P(t|\ell',\ell)||P(t|\bar{\ell},\bar{\ell}')], \tag{53}$$

where the indices in $\sum_{\ell,\ell'\in\mathcal{L}}$ run over the set $\mathcal{L}$ of all visible transitions.

We now focus on the case of a system where two transitions in opposite directions between the same pair of states are visible, $\mathcal{L} = \{+,-\}$, as in single current monitoring. Notice that in this case, the time reversal of a transition is also the visible opposite transition $\bar{\ell} = -\ell$. Thus, the above split of terms simplifies to

$$\sigma_\ell = \langle K\rangle[P(+) - P(-)] \ln \frac{P(+|+)}{P(-|-)} = J_{\mathcal{L}} A_{\text{eff}}, \tag{54}$$

and

$$\begin{aligned}\sigma_t = {} & \langle K\rangle P(+|+)P(+)D[P(t|+,+)||P(t|-,-)] \\ & + \langle K\rangle P(-|-)P(-)D[P(t|-,-)||P(t|+,+)].\end{aligned} \tag{55}$$

The current over the observed transition is $J_{\mathcal{L}} := \langle K\rangle[P(+) - P(-)]$ (cf. Appendix D). In view of the usual bilinear form of the entropy production rate in usual nonequilibrium thermodynamics, we identify the effective affinity $A_{\text{eff}} := \ln P(+|+)/P(-|-)$.

One striking implication of Eq. (51) is that the pairs of alternated transitions $+-$ and $-+$ do not play any role in

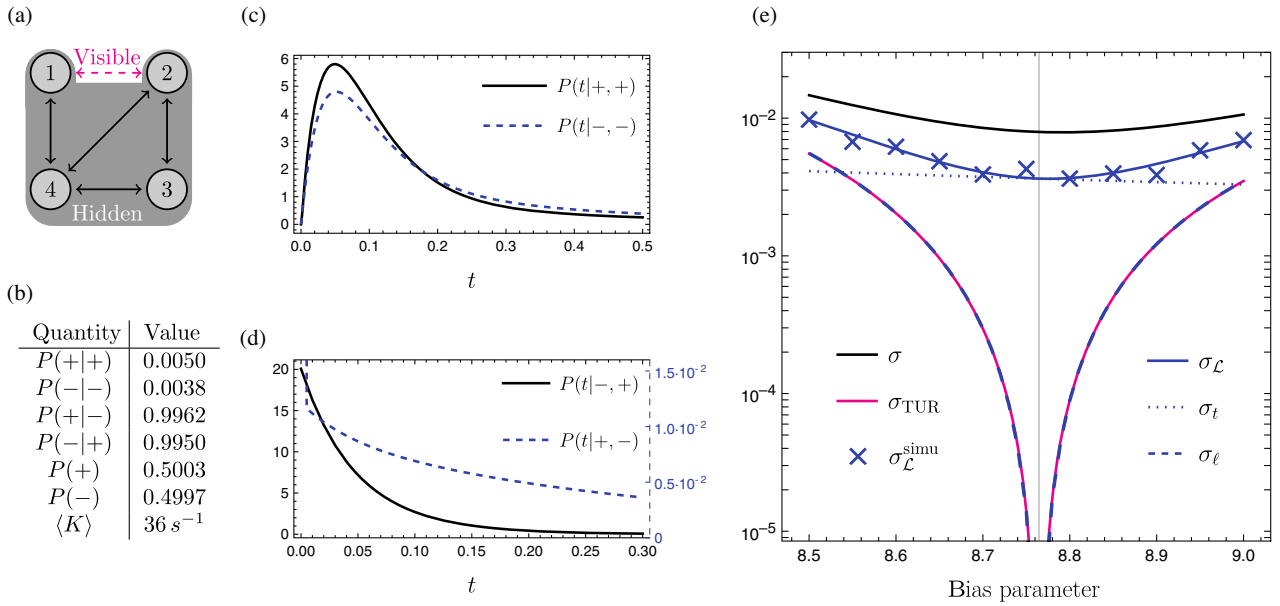


FIG. 6. Estimating the entropy production rate (k_B/s) from repeated transition statistics. (a) Illustration of the inference scheme for a network with the observation of transitions $+ = 1 \to 2$ and $- = 2 \to 1$. From the analytical equations derived in Sec. III B, we show (b) conditional and unconditional probabilities of transitions, and the value of the traffic rate, intertransition-time probability densities for (c) repeated and (d) alternated transitions. In terms of a bias parameter, in panel (e), we show entropy production σ as a solid black curve, the thermodynamic uncertainty relation's lower bound $\sigma_{\rm TUR}$ as a solid magenta curve, and results from Gillespie simulation $\sigma_{\mathcal{L}}^{\rm simu}$ as blue crosses. Next, we show the present results of the inferred entropy production rate $\sigma_{\mathcal{L}}$ in solid blue, its decomposition in intertransition times σ_t is given in dotted blue, and the sequence of transitions σ_ℓ in dashed blue. The vertical line is the value of the (dimensionless) bias parameter for which the visible current vanishes. The transition rates are $W_{24} = W_{34} = W_{41} = W_{43} = 1$, $W_{14} = W_{21} = W_{32} = W_{42} = 20$, and W_{12} equals the exponential of the bias parameter. In panels (b)–(e), the bias parameter is fixed to 8.5. Simulations were performed with a Gillespie algorithm for 2×10^6 s, and the Kullback-Leibler divergence of intertransition times was obtained with an unbiased estimation scheme (cf. main text).

the inferred entropy production. However, the incidence of repeated transitions $++$ and $--$ and their intertransition times contribute to it. This means that only the statistics related to $++$ and $--$ are relevant to irreversibility.

The fact that both σ_ℓ and σ_t are linear combinations of Kullback-Leibler divergences with positive coefficients implies that they are both always equal to or greater than zero. This implies that both $\sigma_\ell \geq 0$ and $\sigma_t \geq 0$ are lower bounds to the rate of entropy production on their own. At equilibrium, $\sigma = \sigma_\ell = \sigma_t = 0$; thus, no irreversibility can be detected from transition frequencies or from intertransition times. Out of equilibrium, however, σ_ℓ and σ_t can vanish in different scenarios, which can be illustrated for the case of observing a single pair of transitions: $\sigma_\ell = 0$ when no net current (computed from frequency of transitions) is found along the visible transition and $\sigma_t = 0$ when the Markov network is unicyclic, i.e., has a ringlike shape, as we show below. In addition, as proved in Ref. [55], if the hidden network either has no cycles or satisfies detailed balance, one also gets $\sigma_t = 0$.

Estimates of entropy production and irreversibility can be extracted from the statistics of single stationary trajectories. A recent example is the thermodynamic uncertainty relation, which allows us to estimate entropy production from empirical time-integrated currents without knowing the transition rates from the bound:

$$\sigma_{\rm TUR} := \frac{2\langle J \rangle^2}{\mathrm{Var}(J)} \leq \sigma, \tag{56}$$

which states that the entropy production rate is lower bounded by the average and variance of any stationary current $J = \lim_{t\to\infty} \langle \sum_{i<j} d_{ij} n_{ij}(t) \rangle / t$ flowing over the system [53,54], with d_{ij} being the asymmetric current increment related to transition $j \to i$ and $n_{ij}(t)$ the number of such transitions in a time interval t. For each trajectory, the stochastic time-integrated current J depends on the number of transitions in each direction. Hence, the full statistics of the sequence of transitions should contain at least the same amount of information as the statistics of J; therefore, we conjecture $\sigma_{\rm TUR} \leq \sigma_\ell$. Furthermore, the intertransition times contribute to the entropy production rate and go unnoticed by $\langle J \rangle$ and $\mathrm{Var}(J)$; therefore, the contribution σ_t contains additional information such as the detection of irreversibility in the absence of net currents.

Figure 6 illustrates how the entropy production inference is obtained using empirical estimates of $P(\pm|\pm)$ and

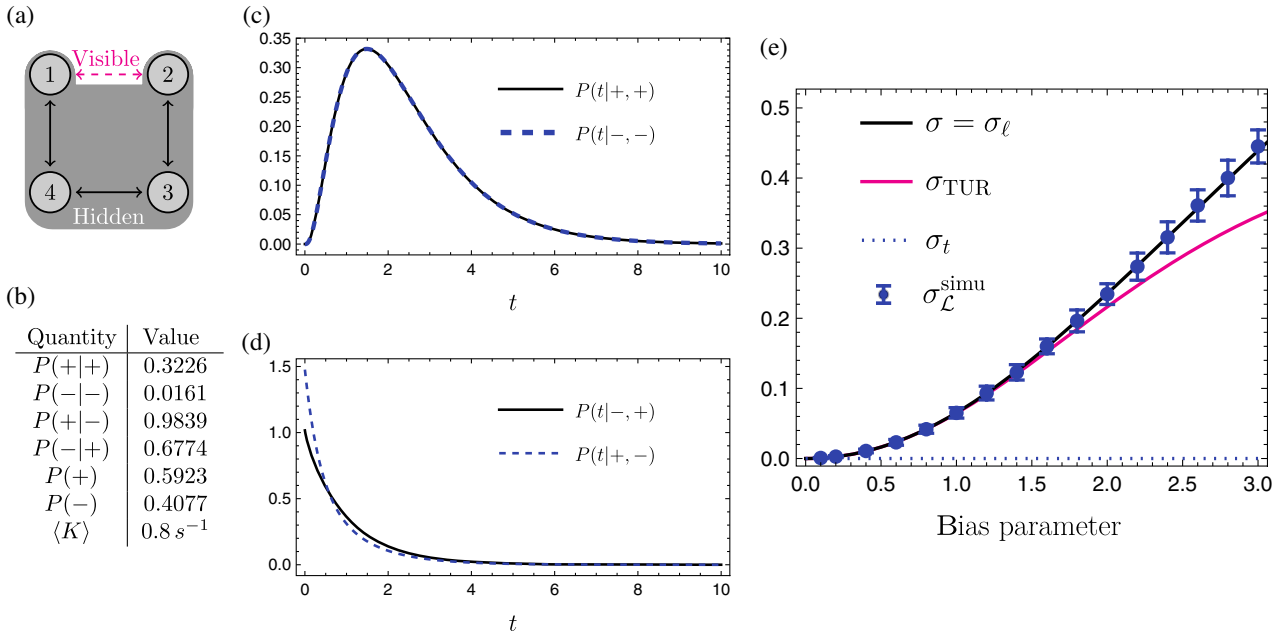


FIG. 7. (a) Ring with four states, ring network, and visible transitions $+ = 1 \to 2$ and $- = 2 \to 1$. (b) Conditional and unconditional probabilities of transitions and the visible traffic rate. (c) Coinciding intertransition-time densities for repeated transitions and (d) for alternated transitions. (e) Summary of the entropy production rate inference scheme. Entropy production σ and the sequence of transition contribution σ_ℓ coincide, both in solid black; the intertransition-time contribution σ_t in dotted blue is shown to vanish; the thermodynamic uncertainty relation $\sigma_{\rm TUR}$ is depicted in solid magenta, and simulations $\sigma_{\mathcal{L}}^{\rm simu}$ in blue dots with error bars. All transition rates are equal to 1 apart from W_{14}, which is the exponential of the dimensionless bias parameter, and entropy production rate dimensions are k_B/s.

$P(t|\pm,\pm)$ as a function of a bias parameter, a value present in transition rates that controls the preference for the performance of a counterclockwise cycle. In a four-state multicyclic network, panel (e) shows the entropy production rate σ (solid black line) that is indeed larger than both σ_ℓ and σ_t. The contribution from the sequence of transitions σ_ℓ (dashed blue) coincides with the thermodynamic uncertainty relation $\sigma_{\rm TUR}$ (solid magenta), and both vanish for a value of bias parameter that stalls the current between $1 \leftrightarrow 2$, which is known as the stalling force. The intertransition time contribution is less sensitive to the bias parameter in this region. It does not vanish at the stalling force, leading to the detection of irreversibility when no net current is visible.

Lastly, for different values of bias parameter, a single trajectory of visible transitions and intertransition times from Gillespie simulations is analyzed in view of Eqs. (54) and (55) to obtain the inferred entropy production rate $\sigma_{\mathcal{L}}^{\rm simu}$ (blue crosses), in good agreement with the analytical $\sigma_{\mathcal{L}}$. Notably, to tackle possible statistical biases that may arise in σ_t from crude histogram-counting procedures, we employ the Pérez-Cruz numerical method [70] that minimizes the statistical bias in the estimation of Kullback-Leibler divergences. See Ref. [71] for our open-source toolbox implementing our estimate of entropy production. Further details of the implementation and convergence analyses are discussed in Appendix E.

B. Ring networks

Networks with ring topology are an important particular case for the inference of irreversibility. They have only one cycle and, therefore, one macroscopic flux and one affinity (thermodynamic force) [68]; such a flux can be obtained from the solution of the master equation, and the affinity is the logarithm of the product of all transition rates $\ln \prod_i (W_{i+1,i}/W_{i,i+1})$. We show that, in this case, the sequence-of-transitions contribution to the inferred entropy production in Eq. (54) provides the exact real entropy production rate, ruling out the necessity of assessing all the microscopic details of stationary probabilities and transition rates.

In this case, the stochastic matrix has a tridiagonal structure plus two terms on its corners, W_{1N} and W_{N1}; without loss of generality, let us consider that $1 \leftrightarrow 2$ is the observed transition, and hence $[\mathbf{S}]_{ij} = W_{ij}(1-\delta_{ij,12})(1-\delta_{ij,21})$. Because of the particular structure of $\mathbf{S}$ in a ring, the Laplace expansion of its inverse leads to the fact that the effective affinity (associated with the visible transitions) in this case equals the cycle affinity A,

$$A_{\rm eff} = \ln \frac{P(+|+)}{P(-|-)} = \ln \prod_i \frac{W_{i+1,i}}{W_{i,i+1}} = A. \tag{57}$$

Analogously, we find that this is also the case for the macroscopic affinity, which can be obtained from the ratio

of conditional transition probabilities or estimated by their respective empirical frequencies.

In this case, $\sigma_\ell = J_{\mathcal{L}}A$, which is the definition of entropy production σ in a cycle. Adding this to the fact proven in Sec. IV that $P(t|+,+) = P(t|-,-)$, implying $\sigma_t = 0$, we find that the inequality between the inferred and real entropy production rates will be saturated and given solely and exactly by the sequence of transitions

$$\sigma_{\mathcal{L}} = \sigma_\ell = \sigma. \tag{58}$$

In other words, the full entropy production of a ring network can be assessed by a single experiment in which a marginal observer collects statistics of the transitions between a single pair of states.

Figure 7 shows the entropy production inference scheme for a ring network of four states. The contribution σ_t vanishes for any value of bias parameter due to the equality of intertransition-time densities for repeated transitions shown in panel (c). The values of σ and σ_ℓ are precisely the same (solid black) as discussed in Eq. (58). Meanwhile, σ_{TUR} provides a lower bound that is approximately saturated for vanishing values of bias parameter, which represents the close-to-equilibrium regime.

VI. INFERENCES FROM VISIBLE TRANSITION IN BIOMOLECULAR SYSTEMS

Here, we apply our theoretical framework to biomolecular machines where partial information, stemming from the observation of a few transitions, is experimentally accessible: for example, DNA polymerase [72], data obtained from single-molecule Förster resonance energy transfer microscopy [23,73], and optical tweezers [33,74] to resolve the displacement of a motor along a track, yet most of the structural and chemical degrees of freedom are hidden. Inspired by these experimental limitations, we first focus on two examples of biologically relevant molecular machines in which we assume that one can only resolve mechanical transitions involving spatial displacements dynein (Sec. VI A) and kinesin (Sec. VI B), which serve as case studies of ring and multicyclic networks, respectively. Next, we extend our study to motors that move in heterogeneous tracks and study the effect of the degree of disorder in the statistics of transitions (Sec. VI D).

A. Dynein ring model

Dyneins are cytoskeletal nanoscale motors that move along microtubules inside cells and perform a varied range of functions, like intracellular cargo transport and beating of flagella [75,76]. Dyneins transduce chemical energy from ATP hydrolysis into mechanical work done by displacing loads along the microtubule.

Here, we study a unicyclic seven-state kinetic model of dynein stepping (cf. Fig. 8), which has a ring topology and

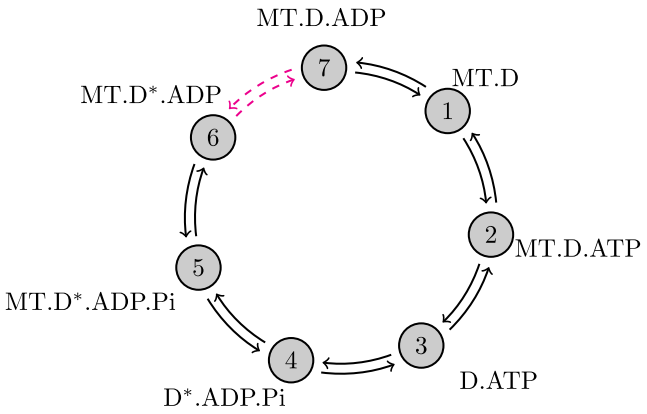


FIG. 8. Sketch of the chemomechanical ring network for dynein with visible transitions $6 \leftrightarrow 7$. The meaning of the transitions between each state and the experimentally inferred values of the transition rates are listed in Table I.

is described in Refs. [77,78]. During every forward-stepping cycle, one ATP molecule binds to the dynein (D) ($1 \to 2$), thereby triggering the release of the dynein from the microtubule (MT) ($2 \to 3$). This is followed by the hydrolysis of ATP that induces a conformational change of the dynein (D^*) ($3 \to 4$) and consequently leads to microtubule binding ($4 \to 5$). In the next step, release of one phosphate group Pi ($5 \to 6$) is followed by a power stroke ($6 \to 7$) and release of one adenosine diphosphate molecule ADP ($7 \to 1$). The different transition rates between these discrete states and their description are listed in Table I.

We consider the setting where single-molecule experiments can follow the cargo displacement and therefore observe only transitions $6 \leftrightarrow 7$. As discussed in Sec. V B, the inferred entropy production rate for this model is

TABLE I. Transition rates for the chemomechanical cycle for the dynein model in Fig. 8 (see Refs. [77,78]). All the rate constants W_{ij} (except W_{12}, W_{32}, and W_{67}, which are in $s^{-1}\,\mu M^{-1}$) are given in units of s^{-1}, and the concentrations in μM. Here, MT refers to the microtubule.

Parameter	Description	Value
W_{17}	ADP release	160
W_{71}	ADP binding	$2.7 \times$ [ADP]
W_{21}	ATP binding	$2 \times$ [ATP]
W_{12}	ATP release	50
W_{32}	MT release in poststroke state	500
W_{23}	MT binding in poststroke state	100
W_{43}	Linker swing to prestroke	1000
W_{34}	Linker swing to poststroke	100
W_{54}	MT binding in prestroke state	10000
W_{45}	MT release in prestroke state	500
W_{65}	Pi release	5000
W_{56}	Pi binding	$0.01 \times$ [Pi]
W_{76}	Power stroke	5000
W_{67}	Reverse stroke	10

TABLE II. Transition rates for the chemomechanical cycle for the kinesin model in Fig. 10 [79]. All the rate constants W_{ij} (except W_{21}, W_{23}, W_{34}, W_{54}, W_{56}, and W_{61}, which are in units of $s^{-1}\,\mu M^{-1}$) are given in units of s^{-1}, and the concentrations in μM.

Parameter	Description	Value
$W^0_{21} = W^0_{54}$	ATP binding	$2.0 \times$ [ATP]
W^0_{12}	Release of ATP	100
$W^0_{32} = W^0_{65}$	ADP release	100
$W^0_{23} = W^0_{56}$	ADP binding	$0.02 \times$ [ADP]
W^0_{25}	ATP binding	0.24
W^0_{52}	Mechanical step	3×10^5
$W^0_{43} = W^0_{16}$	Hydrolysis of ATP	100
$W^0_{34} = W^0_{61}$	Pi binding	$0.02 \times$ [Pi]
W^0_{45}	Release of ATP	$W^0_{12}(W^0_{25}/W^0_{52})^2$

exactly given by $\sigma = \sigma_\ell$ from Eq. (54). From the network topology and transition rates, we evaluate σ_ℓ analytically for different values of parameters such as the concentrations of ATP and ADP. The probabilities of a sequence of two transitions $P(\pm|\pm)$ and $P(\pm|\mp)$ are given from our framework by Eq. (19), and the probability of a single transition is given by Eq. (22).

In Fig. 9, we observe that the entropy production rate increases with the concentration of ATP and decreases with the concentration of ADP. This implies that the forward step of dynein is associated with high dissipation compared to the backward step. The typical dissipation rate for biophysical systems of nanometer to micrometer size ranges between 10 and 1000 k_BT/s [80]. Some examples are kinesin with the dissipation rate 250 k_BT/s and a single RNA hairpin with a dissipation rate between 10 and 250 k_BT/s.

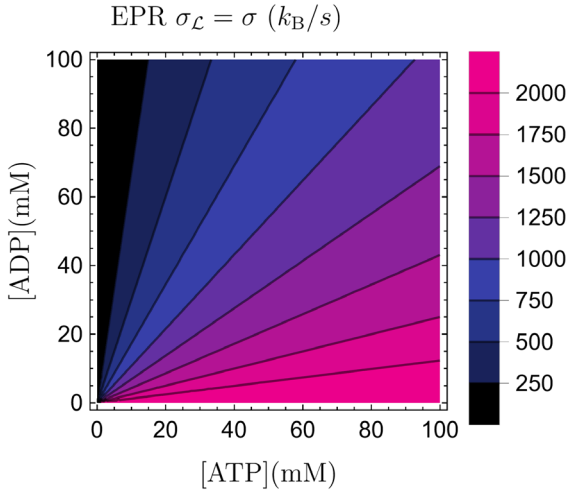


FIG. 9. Entropy production rate in k_B/s for the dynein with visible transitions $6 \leftrightarrow 7$, using rates from Table I and [Pi] = 1 mM, in terms of ATP and ADP.

B. Kinesin multicyclic model

We now study a stochastic model for kinesin motion [79] validated in single-molecule experimental studies [81,82]; see Fig. 10 for an illustration. The model is described by a chemomechanical network comprising six discrete states that describe the mechanism of movement of kinesin on the microtubule. Notice that it has two independent cycles: the "F" cycle $[(1) \to (2) \to (5) \to (6) \to (1)]$ corresponding to the forward motion of kinesin by one step, and the "B" cycle $[(4) \to (5)F \to (2) \to (3) \to (4)]$ resulting in a step backwards. The dynamics along one F cycle is as follows: After ATP binding $(1 \to 2)$, kinesin makes a step forward $(2 \to 5)$ in the filament, followed by ATP hydrolysis that results in the release of one ADP molecule $(5 \to 6)$ and inorganic phosphate Pi $(6 \to 1)$. The backward B cycle proceeds similarly, with the only difference that, after the binding of ATP to kinesin, a backward step along the filament $(5 \to 2)$ occurs. Notice that, in contrast to the model example of dynein, here forward and backward movements are driven by the hydrolysis of one molecule of ATP. The transition rate values are listed in Table VI B. An external load force f biases the transition rates W_{25} and W_{52} involving spatial motion:

$$W_{52}(f) = W^0_{52} e^{-\theta f d_0/k_B T},$$
$$W_{25}(f) = W^0_{25} e^{(1-\theta) f d_0/k_B T}, \tag{59}$$

where θ is the load distribution factor, d_0 is the step size, and f is the load force. On the other hand, for the chemical transitions, we have

$$W_{ij}(f) = 2W^0_{ij}(1 + e^{\chi_{ij} f d_0/k_B T})^{-1}, \tag{60}$$

where χ_{ij} represents the mechanical strain on catalytic domains with $\chi_{ij} = \chi_{ji} > 0$, where $i, j \neq 2, 5$, and the

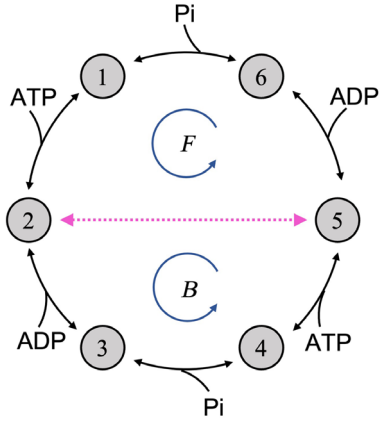


FIG. 10. Sketch of the chemomechanical network model used to describe kinesin motion. The only visible transitions $2 \leftrightarrow 5$ are marked in dotted magenta. Here, F and B denote the cycle corresponding to the forward and backward movement of kinesin, respectively.

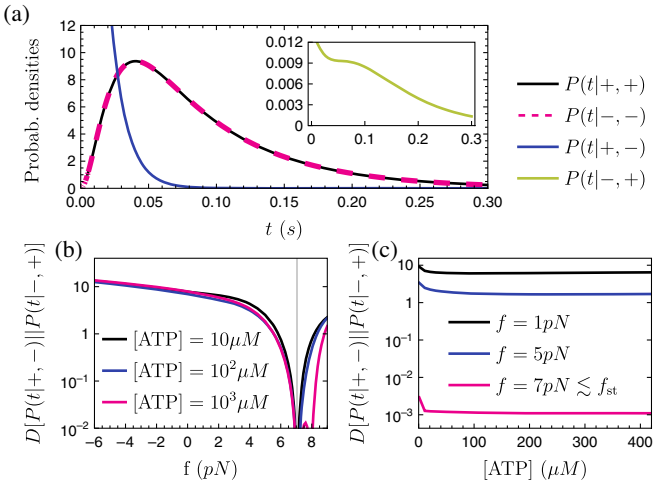


FIG. 11. Exact intertransition-time statistics for the kinesin model in Fig. 10 with visible transitions $2 \leftrightarrow 5$. (a) Intertransition-time densities for every possible pair of transitions. (b) Kullback-Leibler divergence for alternated transitions in terms of the external force f, with the vertical line corresponding to the stalling force of $f_{\mathrm{st}} \sim 7.02$ pN. (c) The same as panel (b), but in terms of the ATP concentration. The rates are displayed Table II, and $[\mathrm{ADP}] = [\mathrm{P}] = 5$ μM.

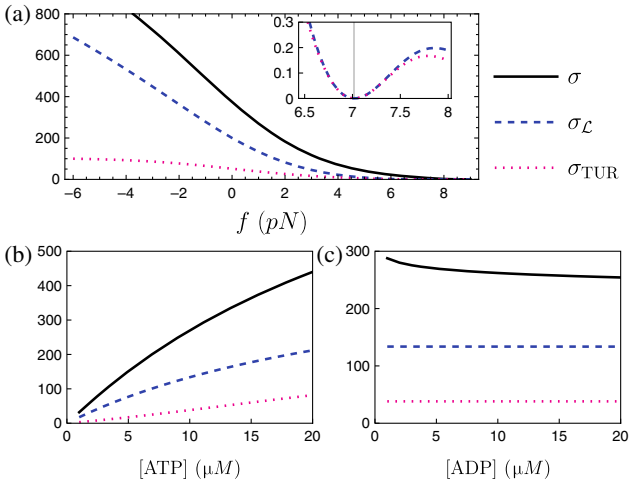


FIG. 12. Exact analytical values for the rate of entropy production for the kinesin model in Fig. 10, with visible transitions $2 \leftrightarrow 5$: entropy production σ of the underlying Markov chain (black curves), inferred entropy production from transition statistics $\sigma_{\mathcal{L}}$ (blue dashed line), and estimate from the thermodynamic uncertainty relation lower bound σ_{TUR} (magenta dotted line). (a) Values in terms of the external load force f and a zoomed-in view around the stalling force in the inset; for this case, $[\mathrm{ATP}] = 10$ μM and $[\mathrm{ADP}] = [P] = 5$ μM. (b) In terms of the concentration of ATP, with $f = 1$ pN and $[\mathrm{ADP}] = [P] = 5$ μM. (c) In terms of the concentration of ADP, with $f = 1pN$, $[\mathrm{ATP}] = 10$ μM, and $[P] = 5$ μM.

concentration of molecular species involved in the chemical transitions are accounted for in most of the W^0_{ij} rates; see Table VI B.

We now focus on the statistics of the transitions associated with the mechanical movement of kinesin, i.e., $2 \leftrightarrow 5$, which are the only ones that can be observed experimentally. In our calculations, we have considered the concentration for $[\mathrm{ADP}] = 70$ μM and $[\mathrm{P}_i] = 1$ mM, the load distribution factor $\theta = 0.65$, $d_0 = 2k_BT$, $\chi_{12} = 0.25$, and $\chi_{56} = \chi_{61} = 0.15$.

Figure 11 shows the intertransition statistics of this model obtained from the analytical expressions in Sec. III, which displays a rich structure due to the multicyclic structure of the model. Our results show that, apart from being defined in a network with two cycles, intertransition-time densities for repeated transitions are identical, $P(t|+,+) = P(t|-,-)$ [Fig. 11(a)]. This property results from the symmetry property that the F and B cycles of the model pass through transitions with identical rates; it is not a generic property for multicyclic networks [see Fig. 6(c)]. As can be seen in Fig. 11(a), the intertransition times have very different densities, due to the transition rates being orders of magnitude apart. In this case, alternated transitions are much faster than repeated ones.

As can be seen in Fig. 11(b), alternated transitions, in general, have different intertransition-time densities, but, for the stall force, $P(t|+,-)$ and $P(t|-,+)$ become similar, as can be seen by minima in $D[P(t|+,-)||P(t|-,+)]$. Both for the force and the concentrations of ATP [cf. Fig. 11(c)], we observe regions of decreasing divergence; however, the entropy production rate is increasing in these regions, which is evidence that intertransition times between alternated transitions do not contribute to the dissipation.

Figure 12 shows the entropy production rate σ and the values inferred from our approach of observing the forward and backward mechanical transition and the thermodynamic uncertainty relation. Figure 12(a) is in terms of the external force f with a zoomed-in view around the stalling force, for which both $\sigma_{\mathcal{L}}$ and σ_{TUR} vanish since there is no flux and no intertransition-time asymmetry between repeated transitions. Figure 12(b) is depicted in terms of the concentration of ATP and (c) that of ADP, from which we observe a monotonic increase of dissipation with [ATP], while it is almost independent of [ADP]. In this model, $\sigma_{\mathcal{L}}$ obtained from Eq. (51), in general, provides a good estimate for σ, generally overperforming the thermodynamic uncertainty relation σ_{TUR}. Because of the absence of σ_t, no dissipation is detected when no net current is present (at the stalling force).

C. Bounds for efficiency of molecular motors

To date, one of the most remarkable applications of the thermodynamic uncertainty relation is the upper bounding of biological motors by the first and second moments of its motion [83,84]. Here, we consider the specific class of molecular motors in which the "stepping transition" does not involve chemical fuel consumption but rather work

done against an external load force, which includes, as specific examples, the dynein and kinesin models of the previous sections. Within this class of molecular motors, the rate of entropy production can be written as $\sigma = (\dot{w}_{\rm chem} - fv)/T$, where $\dot{w}_{\rm chem}$ is the average power exerted on the motor by the chemical transitions (e.g., by the ATP hydrolysis cycle). On the other hand, fv is the average power exerted by the load force f, with v being the net velocity of the motor along the track. For such motors, the second law $\sigma \geq 0$ implies that one can introduce a notion of efficiency as $\eta = fv/\dot{w}_{\rm chem}$, which can be expressed in terms of the entropy production rate as follows:

$$\eta = \frac{1}{1 + T\sigma/fv}. \tag{61}$$

From our observations, we conjecture that the hierarchy of bounds is $\sigma \geq \sigma_{\mathcal{L}} \geq \sigma_{\rm TUR}$, which implies, together with Eq. (61), the following conjectured hierarchy of upper bounds for the molecular motors' efficiency:

$$\eta \leq \frac{1}{1 + T\sigma_{\mathcal{L}}/fv} \leq \frac{1}{1 + T\sigma_{\rm TUR}/fv} \leq 1. \tag{62}$$

Equation (62) implies that the present inference scheme leads to a tighter upper bound to the efficiency than that based on the thermodynamic uncertainty relation introduced in Ref. [83]. Unlike σ_ℓ and $\sigma_{\rm TUR}$, $\sigma_{\mathcal{L}}$ includes information about irreversibility through intertransition times, which shows how the notion of time tightens the efficiency bound.

We illustrate the bounds in Eq. (62) in Fig. 13 for our kinesin model. Its efficiency is positive in the regime where load force and net movement have opposite signs, $0 \leq f \leq f_{\rm st}$ ($f_{\rm st} \approx 7.02$ is the stalling force); thus, the motor performs work against the applied force at the cost of ATP consumption. For the parameter choices that we

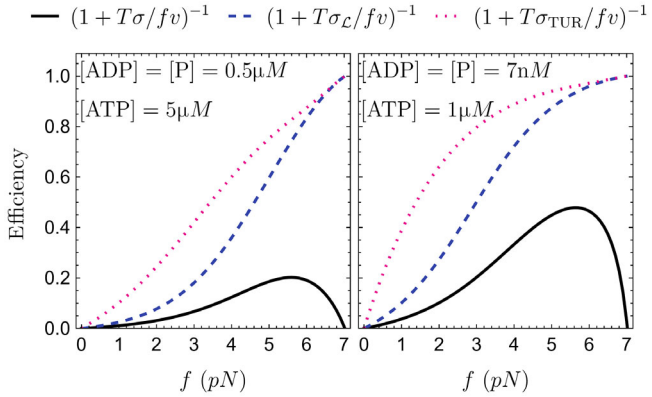


FIG. 13. Efficiency (solid black line) of a kinesin motor doing work against a force f, and its upper bounds obtained from the visible entropy production (dashed blue line) and the thermodynamic uncertainty relation (dotted magenta line). More details are given in Sec. VI B.

explore, we observe that, close to the motor maximum efficiency, the upper bound obtained from transition statistics is 13% closer to the actual value with respect to the estimate obtained from the TUR.

D. Motion on disordered tracks

In many instances, the stochastic motion of molecular machines displays a disordered nature due to the heterogeneity of the track. For example, template-copying machines like DNA and RNA polymerases [85] and ribosomes [86] are often modeled as machines whose motion is dependent on the sequence constituting the track, in such a way that the transition rates depend on the specific monomer type that the machine encounters at every step [87,88]. In this section, we study the effects of the track's disorder in the intertransition statistics associated with the motion of a minimal stochastic model of a molecular machine.

We consider a minimal stochastic model of a molecular machine that moves along a track by burning fuel (i.e., by hydrolysis of ATP). The machine undergoes a series of conformational changes and translocates on a linear heterogeneous track (a polymer) composed of two types of monomers, labeled A and B. We assume that the track is infinite (i.e., we effectively have annealed disorder) and that the generation of the template $q_n \in \{A, B\}$, $n = \{1, 2, \ldots\}$, is an independent and identically distributed process with prescribed probabilities $P(q_n = A) = p$ and $P(q_n = B) = 1 - p$ for the occurrence of A- and B-type monomers, respectively. For our numerical study, we generate the template before running the simulations and use the same template for every run. Figure 14 sketches the disordered nature of a track along the motion of the molecular machine. We also assume that the motor moves following a unicyclic enzymatic reaction composed of four internal configurational states and that only two transitions are visible, $4 \to 1 = +$ and $1 \to 4 = -$, corresponding to forward and backward steps along the track, respectively. The template disorder is implemented in the stochastic model as follows: When the motor reaches a monomer of type q_n, its internal configurational states within one periodicity cell are connected by the rates

$$W_{ij}^{(q_n)} = \begin{cases} W_{ij} & \text{if } q_n = A \text{ and } i, j \neq 4, 1 \\ \alpha W_{ij} & \text{if } q_n = B \text{ and } i, j \neq 4, 1 \\ W_{41} & \text{if } q_{n-1} = A \text{ and } i, j = 4, 1 \\ \alpha W_{41} & \text{if } q_{n-1} = B \text{ and } i, j = 4, 1, \end{cases} \tag{63}$$

where $\alpha \in (0, 1]$ is the disorder factor. This factor scales the transition rates, effectively slowing or accelerating the transitions depending on the track position. As a convention, we have set the transition rate related to a back step $W_{41}^{(q_n)}$ to be defined in terms of the previous monomer's type q_{n-1}. Apart from specific choices of the parameters,

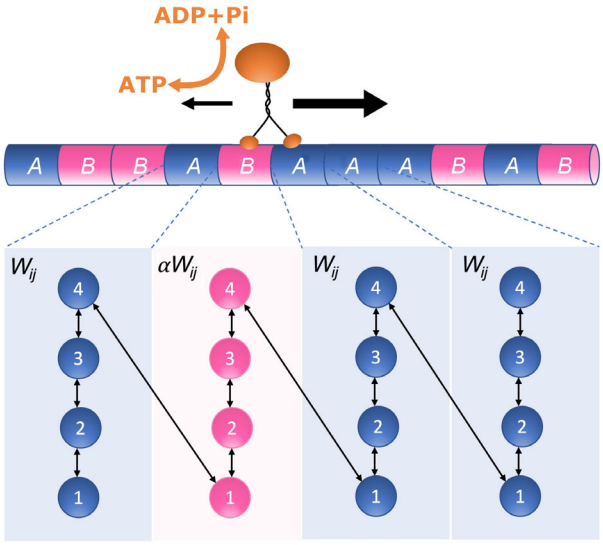


FIG. 14. Sketch of our minimal stochastic model for molecular motor motion in a disordered linear track composed of two types of monomers, A and B. The track is generated as an independent identically distributed sequence of monomers A and B that occur with probabilities p and $1-p$, respectively. Within each periodicity cell of type A or B, the motor internal states have the same structure given by a unicyclic network. The transition rates depend on the motor's position on the track according to Eq. (63), where α is a disorder factor. We also assume that the only visible transitions are those related to translocation to the right (step "forward") $+=4\to 1$ and to the left (step "backward") $-=1\to 4$.

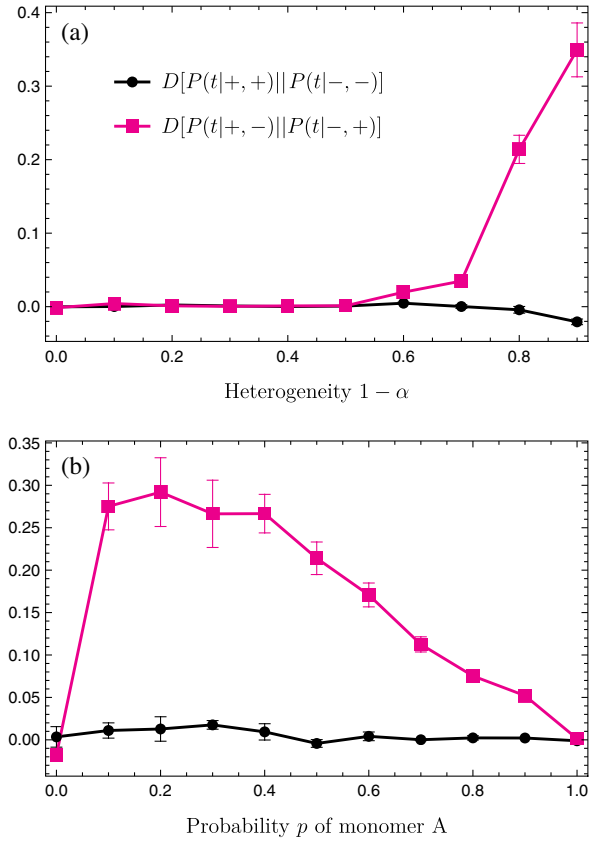


FIG. 15. Kullback-Leibler divergence of the intertransition-time densities for repeated $[D[P(t|+,+)||P(t|-,-)]$, black circles] and alternated $[D[P(t|+,-)||P(t|-,+)]$, magenta squares] transitions obtained from numerical simulations of the model sketched in Fig. 14. We plot the values of these Kullback-Leibler divergences as a function of two parameters of disorder: (a) in terms of heterogeneity $1-\alpha$ with fixed monomer probability $p=0.5$, and (b) in terms of probability p of monomer A, with fixed $\alpha=0.2$. The rates used for Gillespie simulations are $W_{12}=W_{43}=1\ \mathrm{s}^{-1}$, $W_{21}=W_{23}=W_{34}=5\ \mathrm{s}^{-1}$, and $W_{32}=W_{14}=W_{41}=4\ \mathrm{s}^{-1}$, which leads to a nonequilibrium dynamics, and α is introduced according to Eq. (63). Error bars represent the standard deviation from five trajectories, each of duration 5×10^6 s.

this motor has a nonequilibrium dynamics, evidenced by a net drift along the track. In ring topologies, we have observed that intertransition times do not contain irreversibility traces, which is not necessarily true for the disordered case.

We now study how the disorder parameters α and p of this minimal model affects the intertransition-time statistics of successive repeated transitions $++$ and $+-$, and alternated transitions $-+$ and $--$. From the simulation of a molecular motor on a disordered track with four internal states, we observe in Fig. 15 that only the statistics of intertransition times between alternated transitions are affected by the degree of disorder. In particular, we observe, for our example model, that $D[P(t|+,+)||P(t|-,-)]=0$, i.e., a symmetry relation between the intertransition-time distributions of repeated transitions $P(t|+,+)\simeq P(t|-,-)$, which implies $\sigma_t\simeq 0$. As we saw in Sec. V B, such a symmetry relation is a hallmark of unicyclic networks, whereas here we effectively have a multicyclic network with two types of cycles, A and B. We expect the symmetry $P(t|+,+)\simeq P(t|-,-)$, which is already expected for the homogeneous case ($\alpha=1$, $p=0$ or $p=1$), originating from the fact that different monomer types just affect the timescale of the jumps and not the internal network topology within each periodicity cell.

The results for repeated transitions are in stark contrast with our observations for the alternated transitions; see magenta curves in Fig. 15, which shows that $D[P(t|+,-)||P(t|-,+)]$ is strongly dependent on the values of α and p controlling the amount of disorder in the track. Figure 15(a) shows that the degree of asymmetry $D[P(t|+,-)||P(t|-,+)]$ in the alternated intertransition-time statistics increases monotonously with the degree of heterogeneity affecting the timescale of the jumps given by $1-\alpha$. On the other hand, $D[P(t|+,-)||P(t|-,+)]>0$ is also able to probe the presence of sequence heterogeneity as we vary the probability of A monomers p for a fixed α; see Fig. 15(b). Note that, in the latter case, we recover $D[P(t|+,-)||P(t|-,+)]\simeq 0$ for the limiting cases $p=0$

and $p = 1$, which correspond to homogeneous, unicyclic networks. Taken together, these results highlight the possibility of using the intertransition-time statistics between alternated transitions as a probe of the presence of underlying disorder in cyclic enzymatic reactions, which could be further generalized in future work.

VII. DISCUSSION

In this work, we have developed results for generic stationary Markov-jump processes, in and out of equilibrium, whose partial information is restricted to the observation of a partial set among the whole network of transitions. In particular, we investigated the following question: What can one learn from counting the frequency of a partial set of visible transitions and the time elapse between two such visible and successive transitions occurring in a time series? We have tackled the problem of learning dynamic and thermodynamic properties of a system in which only a few transitions are visible to the observer, a novel coarse-graining scheme that proves to be physically meaningful and that provides information through simple relations. For the broad class of stationary Markov processes, we have derived exact analytical results for the conditional and unconditional probability of occurrence of successive transitions and for the time elapsed between successive transitions (intertransition times), which together comprise all the information available to an observer that can only track the occurrence of a few visible transitions.

A key insight of our work is that measuring intertransition times is crucial for thermodynamic inference. Intertransition-time statistics of two successive repeated transitions (e.g., + followed by +) carry different information than that of two successive different transitions (e.g., + followed by −). Repeated transition frequencies and intertransition times contain information about time irreversibility, which can be used to establish tight lower bounds for entropy production even in the absence of probability currents in the transition state space. Counterintuitively, alternated transitions do not contribute to entropy production estimates, but their statistics provide a means to identify the presence of disorder in the hidden state space. Taken together, our work unveils the relevance of intertransition times in thermodynamic inference, putting forward recent works [39,55,89,90] that identified footprints of irreversibility in asymmetries of waiting-time distributions in states rather than in transitions. Exploring symmetry properties and developing inference methods from statistics of a variety of waiting times is a promising novel area of research within the field of stochastic thermodynamics [39,55,89,91–93]. In particular, the companion paper [55] also reports an analysis of waiting-time statistics between transitions and provides complementary results to those developed in our framework and applications.

The results we have obtained are generic and can be applied to large and complex networks, for any given set of visible transitions. One must notice that the inferences become limited when observing a tiny fraction of the transitions on very large networks. Therefore, it would be interesting to study how robust inferences are in relation to the visible portion of the network, in particular, with large and complex structures. In the context of biological systems, this hurdle may be overcome with recent experimental developments. For example, using two-color, single-molecule, photoinduced, electron transfer, fluorescence imaging microscopy [94] and three-color Förster resonance energy transfer [95], one can simultaneously probe multiple conformational changes within an individual biomolecule using one fluorescence color per coordinate. Additionally, the bias in the estimation of relative entropy is circumvented using an unbiased estimator [70], whose implementation is available as an open-source code in Ref. [71]. This open-source toolbox can be used to estimate Kullback-Leibler divergences from experimental time series and, consequently, also the visible entropy production developed herein.

A possible application of the present formalism is to the problem of making insightful considerations about the efficiency of complex biochemical systems or, more in general, of multiterminal systems with more than one input/output [96], or with unknown losses [97]. In fact, while efficiency is well defined when there is one definite input and output, biochemical systems most often involve many sources. For example, in glycolysis, one has ATP, ADP, lactate, water, phosphate, and glucose as metabolites [98]; being the universal energy tokens, it makes sense to consider the ratio of ADP to ATP production as a measure of efficiency, but then the problem is how to single them out of all other mechanisms and make claims about the efficiency of the process. Furthermore, in more complex biochemical networks, such as those that also involve respiration, one might want to focus on other metabolites (oxygen, carbon dioxide, etc.). To develop such an approach, it is thus mandatory to develop a more phenomenological theory that is consistent with the fundamental tenets of thermodynamics but can also be adapted to the specific tasks or instruments that the observer has in mind. In this respect, the theory presented here may provide a general conceptual and operational scheme.

We illustrated our results in two models of motion of molecular motors that have been validated with experimental data, revealing that our methodology could be applied to real data extracted from, e.g., single-molecule experiments. We expect that our generic inference techniques will be applied to other disciplines where partially observed transitions emerge, such as diagnosis algorithms [99], finite automata [8,100], Markov decision processes [101], disease spreading [102], information machines [103], probing of open (quantum) systems [104,105], and Maxwell demons [106].

ACKNOWLEDGMENTS

We are thankful to Ken Sekimoto and Nahuel Freitas for useful discussions. We also thank Fahad Kamulegeya for preliminary numerical results. P. E. H. acknowledges support from Grants No. 2017/24567-0 and No. 2020/03708-8, São Paulo Research Foundation (FAPESP), and Massimiliano Esposito for hosting. P. E. H., and A. D. acknowledge financial support from the ICTP Quantitative Life Sciences section. M. P. acknowledges the National Research Fund Luxembourg (Project CORE ThermoComp C17/MS/11696700) and the European Research Council, Project NanoThermo (ERC-2015-CoG Agreement No. 681456) for support.

APPENDIX A: PROOF OF EQ. (18)

We now prove Eq. (18) in the main text as follows. We map the first-transition time problem into a first-passage time problem by introducing auxiliary absorbing states for each transition in $\mathcal{L}$. This procedure is inspired by recent work on first-passage times between states in a Markov chain [49] and is also similar to the manipulation of networks to obtain current statistics by creating copies of some states introduced by Hill [107,108]. We use the fact that the survival probability density of a process described by stochastic matrix $\mathbf{W}$ and starting in state i does not reach state j by time t and is given by [48,49]

$$\mathbb{S}(t,j|i) = \sum_{k\neq j}\langle k|\exp(t\mathbf{W})|i\rangle = 1 - \langle j|\exp(t\mathbf{W})|i\rangle, \quad \text{(A1)}$$

and rewrite this result for transitions rather than states.

We consider a continuous-time Markov jump process over an irreducible network of discrete states $\Omega = \{1,2,\ldots,N\}$. We introduce auxiliary absorbing states s_i (sinks) for $i \in [1,|\mathcal{L}|]$ to account for the occurrence of every transition $\ell \in \mathcal{L}$ separately. The $|\Omega_{\rm ex}| \times |\Omega_{\rm ex}|$ stochastic matrix $\mathbf{W}_{\rm ex}$ associated with the dynamics over the extended state space $\Omega_{\rm ex} \coloneqq \Omega \cup \{s_1,\ldots,s_{|\mathcal{L}|}\}$ is such that every element of the visible set $\ell \in \mathcal{L}$, a visible transition, is redirected to point towards its associated sink s_ℓ (cf. Fig. 16), and since the sink is an absorbing state, we set $[\mathbf{W}_{\rm ex}]_{j,s_i} = 0$ for all j in Ω. Following our notation, the sources of visible transitions $\langle\langle \ell||$ are preserved while the targets are redirected to the respective sinks $||\ell\rangle\rangle \to |s_\ell\rangle$.

The extended matrix $\mathbf{W}_{\rm ex}$ has four blocks; the top-left block is the survival matrix $\mathbf{S}$ with size $|\Omega| \times |\Omega|$, and both blocks to the right are zero matrices. The bottom-left block $\mathbf{L}$ has size $|\mathcal{L}| \times |\Omega|$ and contains the redirected transitions; mathematically, it is expressed as $\mathbf{L} = \sum_{j=1}^{|\mathcal{L}|}\langle\langle \mathcal{L}_j||\mathbf{W}^{\mathsf{T}}||\mathcal{L}_j\rangle\rangle|j\rangle\langle\langle \mathcal{L}_j||$, where the sum $\sum_{j=1}^{|\mathcal{L}|}$ runs through every element of the visible set of transitions. For the example in Fig. 16, the extended stochastic matrix is

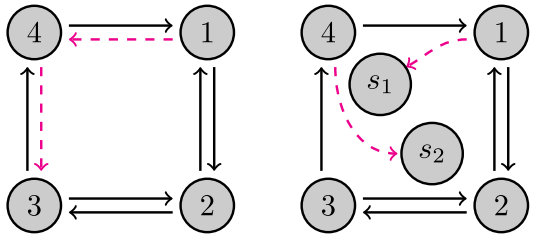


FIG. 16. Left diagram: example of a four-states network with visible transitions $\mathcal{L} = \{1 \to 4, 4 \to 3\}$ (dashed magenta line). Right diagram: network of the extended state space, where visible transitions are redirected to auxiliary absorbing states s_1 and s_2.

$$\mathbf{W}_{\rm ex} = \begin{array}{c} \begin{array}{cccc|cc} 1 & 2 & 3 & 4 & s_1 & s_2 \end{array} \\ \left(\begin{array}{cccc|cc} W_{11} & W_{12} & W_{13} & W_{14} & 0 & 0 \\ W_{21} & W_{22} & W_{23} & W_{24} & 0 & 0 \\ W_{31} & W_{32} & W_{33} & 0 & 0 & 0 \\ 0 & W_{42} & W_{43} & W_{44} & 0 & 0 \\ W_{41} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & W_{34} & 0 & 0 \end{array}\right) \end{array}. \quad \text{(A2)}$$

Since the last columns are zero, any power $n \geq 1$ of the matrix has the property that its left blocks only depend on powers of themselves, and the right blocks remain zero,

$$\begin{aligned}[\mathbf{W}_{\rm ex}^n]_{i,j} &= \sum_{k\in\Omega}[\mathbf{W}_{\rm ex}]_{i,k}[\mathbf{W}_{\rm ex}^{n-1}]_{k,j} \\ &\quad + \sum_{k'=1}^{|\mathcal{L}|}[\mathbf{W}_{\rm ex}]_{i,\Omega+k'}[\mathbf{W}_{\rm ex}^{n-1}]_{\Omega+k',j} \\ &= \sum_{k\in\Omega}[\mathbf{W}_{\rm ex}]_{i,k}[\mathbf{W}_{\rm ex}^{n-1}]_{k,j}, \end{aligned} \quad \text{(A3)}$$

where the second equality follows from $[\mathbf{W}_{\rm ex}]_{i,\Omega+k'} = 0$ for any i and $1 \leq k' \leq |\mathcal{L}|$. Hence, its matrix exponential $\exp(\mathbf{W}_{\rm ex}) = \sum_{k=0}^{\infty}\mathbf{W}_{\rm ex}^k/k!$ is such that its top-left block is its own exponential, i.e.,

$$\exp(t\mathbf{W}_{\rm ex}) = \left(\begin{array}{c|c} \exp(t\mathbf{S}) & \mathbf{0}_{|\Omega|\times|\mathcal{L}|} \\ \hline \mathbf{L}\mathbf{S}^{-1}(\exp(t\mathbf{S}) - \mathbf{1}) & \mathbf{1}_{|\mathcal{L}|\times|\mathcal{L}|} \end{array}\right), \quad \text{(A4)}$$

where the last column has a nonsquare block of zeros and one $|\mathcal{L}| \times |\mathcal{L}|$ identity matrix.

We are interested in the case where a visible transition was performed at time zero; thus, the initial state is $||\ell_i\rangle\rangle$, $\ell_i \in \mathcal{L}$, and the next visible transition $\ell_{i+1} \in \mathcal{L}$ is performed by time t, which is equivalent to the first-passage distribution. Hence, the transition analogue of Eq. (A1) is the survival density related to the respective sink $s_{\ell_{i+1}}$:

$$\mathbb{S}(t, s_{\ell_{i+1}}|\ell_i) = 1 - \langle s_{\ell_{i+1}}|\exp(t\mathbf{W}_{\rm ex})||\ell_i\rangle\rangle = 1 - \langle s_{\ell_{i+1}}|\mathbf{L}\mathbf{S}^{-1}(\exp(t\mathbf{S}) - \mathbf{1})||\ell_i\rangle\rangle, \quad \text{(A5)}$$

where the last equality comes from the fact that the matrix element in question belongs to the bottom-left block of Eq. (A4). The first-transition distribution is given by the time derivative $\mathbb{F} = -\partial_t\mathbb{S}$; hence,

$$\begin{aligned}
&\mathbb{F}(t, s_{\ell_{i+1}}|\ell_i)\\
&= -\partial_t\mathbb{S}(t, s_{\ell_{i+1}}|\ell_i)\\
&= \partial_t\langle s_{\ell_{i+1}}|\mathbf{L}\mathbf{S}^{-1}(\exp(t\mathbf{S}) - \mathbf{1})||\ell_i\rangle\rangle\\
&= \langle s_{\ell_{i+1}}|\mathbf{L}\exp(t\mathbf{S})||\ell_i\rangle\rangle\\
&= \langle s_{\ell_{i+1}}|\left(\sum_{j=1}^{|\mathcal{L}|}\langle\langle\mathcal{L}_j||\mathbf{W}^{\mathsf{T}}||\mathcal{L}_j\rangle\rangle|j\rangle\langle\langle\mathcal{L}_j||\right)\exp(t\mathbf{S})||\ell_i\rangle\rangle\\
&= \langle\langle\ell_{i+1}||\mathbf{W}^{\mathsf{T}}||\ell_{i+1}\rangle\rangle\langle\langle\ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle, \quad \text{(A6)}
\end{aligned}$$

which provides the desired result

$$P(t, \ell_{i+1}|\ell_i) = \langle\langle\ell_{i+1}||\mathbf{W}^{\mathsf{T}}||\ell_{i+1}\rangle\rangle\langle\langle\ell_{i+1}||\exp(t\mathbf{S})||\ell_i\rangle\rangle \quad \text{(A7)}$$

for the joint probability density that a transition ℓ_{i+1} happens at a time t given that ℓ_i was performed at time zero and no other visible transition happened in between. ∎

APPENDIX B: COMBINATORICS OF RING NETWORKS

For a ring topology with reversible edges as described in Sec. IV, we need to evaluate the coefficient $C^{\vec{k}}_{+,+}$. We observe the following:

(i) There are $\prod_{i=1}^{N-1}(k_i + 1)$ possible backbones since there are $k_i + 1$ clockwise edges between states i and $i + 1$. When the dynamics is occurring, the chosen backbone is now a set of prohibited edges that are saved for last in order to make sure that the cycle is completed.

(ii) When the system is at state $1 < i < N$ for the first time, it can choose between k_i clockwise and k_{i-1} counterclockwise edges. Next time i is visited, there will be one less way out of it since one transition was already used, and so on. Therefore, we have the contribution $(k_i + k_{i-1})!$ from each state.

(iii) The contributions of states 1 and N are, respectively, $k_1!$ and $k_{N-1}!$ since from there it is only possible to jump in one direction.

(iv) Edges with the same direction and connecting the same pair of states are indistinguishable, so we have to divide everything by their number of permutations $(k_i + 1)!k_i!$.

By gathering every contribution discussed above, we obtain

$$C^{\vec{k}}_{+,+} = \prod_{i=2}^{N-1}\binom{k_i + k_{i-1}}{k_i}. \quad \text{(B1)}$$

For the alternated case, notice the following:

(i) There are $\prod_{i=1}^{M} k_i$ possible backbones.

(ii) When the system is at state $1 < n < M + 1$ for the first time, it can choose between k_i clockwise and $k_{i-1} - 1$ counterclockwise edges (one is from the backbone). Next time i is visited, there will be one less way out of it since one transition was already used. Therefore, we have the contribution $(k_i + k_{i-1} - 1)!$ of each state.

(iii) The contributions of states 1 and $M + 1$ are, respectively, $k_1!$ and $(k_M - 1)!$.

(iv) Because of the indistinguishability of edges, everything is divided by $k_n!^2$.

Thus, we have the coefficient

$$C^{\vec{k},M}_{+,-} = \prod_{i=2}^{M}\binom{k_i + k_{i-1} - 1}{k_i}. \quad \text{(B2)}$$

APPENDIX C: DIAGRAMMATIC APPROACH

The division $\hat{P}(s|+,+)/\hat{P}(0|+,+)$ involves a product of $\hat{\pi}_{i-1,i}(s)/(\hat{\pi}_{i-1,i}(0)/$, which, by definition, is $W_{ii}/(W_{ii}+s)$. Therefore,

$$\frac{\hat{P}(s|+,+)}{\hat{P}(0|+,+)} = \left(\prod_{i=1}^{N}\frac{W_{ii}}{W_{ii}+s}\right)\frac{[\prod_{i=1}^{N-1}1 - \Theta[x_i]]_{s=0}}{\prod_{i=1}^{N-1}1 - \Theta[x_i]} \quad \text{(C1)}$$

and

$$\frac{\hat{P}(s|-,-)}{\hat{P}(0|-,-)} = \left(\prod_{i=1}^{N}\frac{W_{ii}}{W_{ii}+s}\right)\frac{[\prod_{i=1}^{N-1}1 - \Xi(x_i)]_{s=0}}{\prod_{i=1}^{N-1}1 - \Xi(x_i)}. \quad \text{(C2)}$$

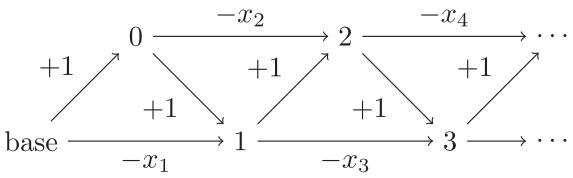
The first factors in the right-hand side of both equations above are the same, and using a diagrammatic approach, we show that the second one is also the same.

Recall that the continued fraction generators used in Sec. IV are defined as $\Theta[x_{i+1}] = x_{i+1}/(1 - \Theta[x_i])$, $\Theta[x_1] = x_1$, which generates a continued fraction. To evaluate the monomial $\prod_i(1 - \Theta[x_i])$, we notice that if we pick two consecutive terms, they simplify to

$$(1 - \Theta[x_{i+1}])(1 - \Theta[x_i]) = 1 - \Theta[x_i] - x_{i+1}, \quad \text{(C3)}$$

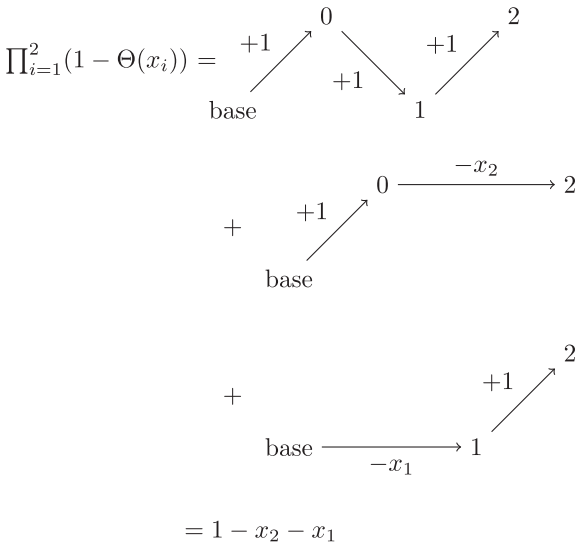
which means that each multiplication of consecutive terms will lead to two terms: one that is only $(1 - \Theta[x_i])$ and the other $-x_{i+1}$. In other words, there are two possible paths that will be added up to evaluate the whole product.

This branching procedure can be portrayed by the diagram below, where each node j represents the value of $\prod_{i=1}^{j}(1-\Theta[x_i])$. Notice that at each transition, there are two arrows arriving, one representing the multiplication by $+1$ and the other by $-x_j$, as discussed.



The value of the monomial at each transition can be evaluated as the sum of all paths starting from the base and reaching such a transition, weighted by the product of the weight of all edges involved.

To illustrate this case, consider node 2:



For the reversed fraction generator $\Xi[x_i]$, the analogue of Eq. (C3) is

$$(1-\Xi[x_{i+1}])(1-\Xi[x_i]) = 1-\Xi[x_{i+1}] - x_i, \quad \text{(C4)}$$

which means that the diagram has the same structure and, more importantly, it is covered in the backwards direction. The arrows point in the other direction, but the weights remain unchanged, leading to the final result when $\prod_i(1-\Theta[x_i])$ and $\prod_i(1-\Xi[x_i])$ start and end at the same points, which implies

$$\prod_{i=1}^{N-1}(1-\Theta[x_i]) = \prod_{i=1}^{N-1}(1-\Xi[x_i]). \quad \text{(C5)}$$

This property guarantees the Haldane-like equality Eq. (33).

APPENDIX D: IRREVERSIBILITY IN TRANSITION TIME SERIES

We now analytically evaluate the rate of irreversibility in the time series $\Gamma_t^{\mathcal{L}}$, which is defined by Eq. (46), copied here for convenience:

$$\sigma_{\mathcal{L}} = \lim_{\tau\to\infty}\frac{1}{\tau}D(P[\Gamma_\tau^{\mathcal{L}}]||P[\bar{\Gamma}_\tau^{\mathcal{L}}]), \quad \text{(D1)}$$

where the probability of a trajectory can be expressed as the joint probability of all random variables involved, the visible transitions, and intertransition times

$$P[\Gamma_\tau^{\mathcal{L}}] = P(t_0,\ell_0,t_1,\ell_1,\ldots) \equiv P(\vec{t},\vec{\ell}). \quad \text{(D2)}$$

With no further assumptions, the path probability of the sequence of transitions can be cast as the product of the sequence probability and intertransition times as $P[\Gamma_\tau^{\mathcal{L}}] = P(\vec{\ell})P(\vec{t}|\vec{\ell})$. Since the underlying process is Markovian and there are no transitions with the same source and target, the sequence of transitions is also Markovian, $P(\ell_i|\ell_{i-1},\ldots,\ell_0) = P(\ell_i|\ell_{i-1})$; thus, the probability of a sequence of transitions in a trajectory is

$$P(\vec{\ell}) = P(\ell_0)\prod_{i=1}^{n}P(\ell_i|\ell_{i-1}), \quad \text{(D3)}$$

where n is the total number of transitions within $\Gamma_\tau^{\mathcal{L}}$. On the other hand, the probability of a sequence of intertransition times conditioned to the occurrence of a given sequence of transitions reads

$$P(\vec{t}|\vec{\ell}) = \prod_{i=1}^{n}P(t_i|\ell_{i-1},\ell_i). \quad \text{(D4)}$$

Therefore, the inferred entropy production rate reads

$$\begin{aligned}\sigma_{\mathcal{L}} &= \lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\Gamma_\tau^{\mathcal{L}}}P[\Gamma_\tau^{\mathcal{L}}]\ln\frac{P[\Gamma_\tau^{\mathcal{L}}]}{P[\bar{\Gamma}_\tau]} \\ &= \lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\vec{\ell}}\int \mathrm{d}t_0\cdots\mathrm{d}t_n \\ &\quad\times\left\{P[\Gamma]\ln\frac{P[\vec{\ell}]}{P[\bar{\vec{\ell}}]} + P[\Gamma]\ln\frac{P[\vec{t}|\vec{\ell}]}{P[\bar{\vec{t}}|\bar{\vec{\ell}}]}\right\}. \quad \text{(D5)}\end{aligned}$$

The first term in Eq. (D5) is

$$\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\vec{\ell}}\int \mathrm{d}\vec{t}P[\Gamma_\tau^{\mathcal{L}}]\ln\frac{P(\ell_0)P(\ell_1|\ell_0)\cdots}{P(\bar{\ell}_n)P(\bar{\ell}_{n-1}|\bar{\ell}_n)\cdots}$$
$$=\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\vec{\ell}}P[\vec{\ell}]\left\{\ln P(\ell_0)+\ln\frac{P(\ell_1|\ell_0)}{P(\bar{\ell}_0|\bar{\ell}_1)}+\ldots\right\}$$
$$=\langle K\rangle\sum_{\ell,\ell'\in\mathcal{L}}P(\ell|\ell')P(\ell')\ln\frac{P(\ell|\ell')}{P(\overline{\ell'}|\bar{\ell})}$$
$$=:\sigma_\ell. \tag{D6}$$

The second term in Eq. (D5) reads

$$\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\vec{\ell}}\int \mathrm{d}\vec{t}P[\Gamma_\tau^{\mathcal{L}}]\ln\frac{P(t_1|\ell_0,\ell_1)\cdots}{P(t_n|\bar{\ell}_n,\bar{\ell}_{n-1})\cdots}$$
$$=\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{\vec{\ell}}P[\vec{\ell}]\left\{\int \mathrm{d}t_1P(t_1|\ell_0,\ell_1)\ln\frac{P(t_1|\ell_0,\ell_1)}{P(t_1|\bar{\ell}_1,\bar{\ell}_0)}+\ldots\right\}$$
$$=\langle K\rangle\sum_{\ell,\ell'\in\mathcal{L}}P(\ell|\ell')P(\ell')D[P(t|\ell',\ell)||P(t|\bar{\ell},\overline{\ell'})]$$
$$=:\sigma_t. \tag{D7}$$

For the special case of a single visible transition that can only take values $\mathcal{L}=\{+,-\}$, the time reversal of $+$ is $-$ and vice versa. Inference of the entropy production rate's first term simplifies to

$$\sigma_\ell=\langle K\rangle[P(+|+)P(+)-P(-|-)P(-)]\ln\frac{P(+|+)}{P(-|-)}. \tag{D8}$$

The stationary occupation probability vector can be found by

$$p_\infty(j)=(-1)^{i+j}\det(\mathbf{W}_{\backslash(i,j)}),\quad \forall\ i \tag{D9}$$

since $\det(\mathbf{W})=\sum_i W_{ij}(-1)^{i+j}\det(\mathbf{W}_{\backslash(i,j)})=0$ and $\sum_j W_{ij}p_\infty(j)=0$ for every transition matrix defining a Markov chain. Lastly, by the construction of the survival matrix $\mathbf{S}$, we observe that $\mathbf{S}_{\backslash(i,i)}=\mathbf{W}_{\backslash(i,i)}$ and $\mathbf{S}_{\backslash(j,j)}=\mathbf{W}_{\backslash(j,j)}$. Now, we are in the position to show that

$$\begin{aligned}P(+|+)&=1-P(-|+)\\&=1-\langle\!\langle -||\mathbf{W}^{\mathsf{T}}||-\rangle\!\rangle\langle\!\langle -||\mathbf{S}^{-1}||+\rangle\!\rangle\\&=1-\langle\!\langle -||\mathbf{W}^{\mathsf{T}}||-\rangle\!\rangle\frac{\det(\mathbf{S}_{\backslash(2,2)})}{\det(\mathbf{S})}\\&=1-\langle\!\langle -||\mathbf{W}^{\mathsf{T}}||-\rangle\!\rangle\frac{\det(\mathbf{W}_{\backslash(2,2)})}{\det(\mathbf{S})}\\&=1-\langle\!\langle -||\mathbf{W}^{\mathsf{T}}||-\rangle\!\rangle\frac{\langle\!\langle -||p_\infty\rangle}{\det(\mathbf{S})}\\&=1-\frac{\langle K\rangle P(-)}{\det(\mathbf{S})},\end{aligned} \tag{D10}$$

and analogously, we obtain $P(-|-)=1-\langle K\rangle P(+)/\det(\mathbf{S})$. Plugging this latter equation into Eq. (D8), it simplifies to

$$\langle K\rangle[P(+)-P(-)]\ln\frac{P(+|+)}{P(-|-)}=:J_{\mathcal{L}}A_{\mathrm{eff}}, \tag{D11}$$

which is Eq. (54) in the main text. The factor $\langle K\rangle[P(+)-P(-)]$ is the definition of the flux through the observed transition $J_{\mathcal{L}}$, suggesting the definition of the second factor as the effective affinity A_{eff}.

Following the same reasoning, the intertransition-time contribution simplifies to

$$\begin{aligned}\sigma_t=&\langle K\rangle P(+|+)P(+)D[P(t|+,+)||P(t|-,-)]\\&+\langle K\rangle P(-|-)P(-)D[P(t|-,-)||P(t|+,+)],\end{aligned} \tag{D12}$$

which is Eq. (55) in the main text. The sum of Eqs. (D8) and (D12) results in the entropy production rate inferred by an observer who only accesses two opposite transitions between a single pair of states.

APPENDIX E: KULLBACK-LEIBLER DIVERGENCE FROM FINITE DATA

Estimation of Kullback-Leibler divergences between distributions of continuous random variables, such as that present in σ_t, from time series is not a straightforward task, as it can lead to systematic errors and statistical biases [44,109,110]. Furthermore, inference schemes to deal with finite data have been largely explored in the analytical sense, as discussed in this paper, hence the need for accurate estimators.

The most intuitive approach involves estimating the probability distributions (here, P and Q) via standard histogram counting methods of the data collected from an experiment or simulation, and later approximating the integral $D[P(x)||Q(x)]=\int \mathrm{d}xP(x)\ln P(x)/Q(x)\simeq\sum_i \mathsf{P}_i\ln(\mathsf{P}_i/\mathsf{Q}_i)$, with P_i and Q_i the probability for the data to fall in the ith bin. However, this approach leads to a biased estimate of the Kullback-Leibler divergence, as shown in previous work [44,109,110]. A method developed in Ref. [70] explores an alternative, unbiased estimation method for the bias-free Kullback-Leibler divergence, which is based on the comparison between the cumulative distributions of two independent data sets generated by P and Q. This method was adapted to the estimate of the intertransition-time Kullback-Leibler divergences shown in the analysis of simulated results throughout the main text. We made our code open source and available in Ref. [71], with further details and illustrations of generating visible transitions' time series, and evaluating Kullback-Leibler divergences and $\sigma_{\mathcal{L}}$.

Briefly, the method consists of taking two finite data sets sampled from two independent processes with distributions

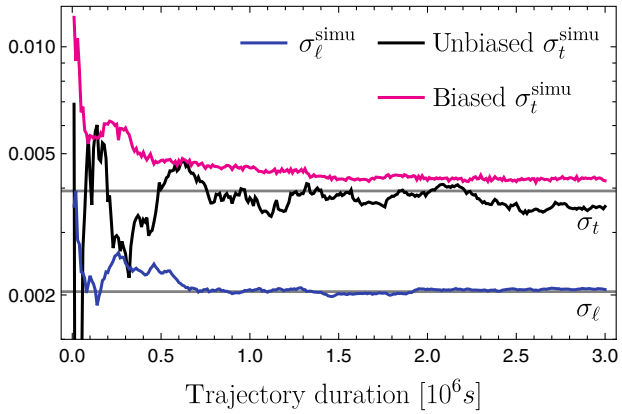


FIG. 17. Convergence analysis of the two contributions to the entropy production rate inferred from numerical simulations $\sigma_{\mathcal{L}}^{\rm simu}$ for the model shown in Fig. 6 with bias parameter 8.6. The magenta line (top) is the intertransition times' Kullback-Leibler divergence $\sigma_t^{\rm simu}$ evaluated by the biased method from histogram counting. The black line (middle) is the intertransition times' Kullback-Leibler divergence $\sigma_t^{\rm simu}$ obtained using the unbiased method [70,71]. The blue line (bottom) is the estimate of the Kullback-Leibler divergence from the transitions' occurrence statistics from $\sigma_\ell^{\rm simu}$. The analytical values of both σ_t and σ_ℓ are shown by horizontal gray lines.

$P(x)$ and $Q(x)$. Linear interpolations $F_c(x)$ and $F_c(x)$ of their associated empirical cumulative distributions are obtained for small enough ϵ, which are used in the following estimate shown to converge to the Kullback-Leibler divergence in Ref. [70],

$$\lim_{n\to\infty}\left[\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{F_c(X_i)-F_c(X_i-\epsilon)}{F_c(X_i)-F_c(X_i-\epsilon)}\right)-1\right] = D[P(x)||Q(x)], \tag{E1}$$

where n is the number of points X_i from the data with distribution $P(x)$.

Figure 17 shows the convergence of σ_ℓ and σ_t estimates as the number of data points in the time series increases. The value of $\sigma_\ell^{\rm simu}$ has a fast convergence using the empirical frequencies of transitions. For σ_t, however, the estimate depends strongly on the method used as intertransition times are continuous random variables. By estimating the probability distributions with kernel density estimations methods and numerically evaluating the Kullback-Leibler divergence's integral, we obtain $\sigma_t^{\rm simu}$ (magenta line), which leads to a significant bias above the expected value. On the other hand, using Eq. (E1) through the resources in Ref. [71], the estimate $\sigma_t^{\rm simu}$ (black line) displays no evident sign of statistical bias above or below the analytical value of σ_t. This result motivated the usage of Eq. (E1) as our estimate for all the Kullback-Leibler divergences in this work.

[1] Nicolaas Godfried Van Kampen, *Stochastic Processes in Physics and Chemistry*, Vol. 1 (Elsevier, New York, 1992).
[2] Abraham Tamir, *Applications of Markov Chains in Chemical Engineering* (Elsevier, New York, 1998).
[3] David F Anderson and Thomas G Kurtz, *Continuous Time Markov Chain Models for Chemical Reaction Networks*, in *Design and Analysis of Biomolecular Circuits* (Springer, New York, 2011) pp. 3–42.
[4] Francesco Avanzini, Emanuele Penocchio, Gianmaria Falasco, and Massimiliano Esposito, *Nonequilibrium Thermodynamics of Non-Ideal Chemical Reaction Networks*, J. Chem. Phys. **154**, 094114 (2021).
[5] Linda J. S. Allen, *An Introduction to Stochastic Processes with Applications to Biology* (CRC Press, Boca Raton, 2010).
[6] Anatoly B. Kolomeisky and Michael E. Fisher, *Molecular Motors: A Theorist's Perspective*, Annu. Rev. Phys. Chem. **58**, 675 (2007).
[7] Debashish Chowdhury, *Modeling Stochastic Kinetics of Molecular Machines at Multiple Levels: From Molecules to Modules*, Biophys. J. **104**, 2331 (2013).
[8] David H. Wolpert, *The Stochastic Thermodynamics of Computation*, J. Phys. A **52**, 193001 (2019).
[9] Hao Ge and Hong Qian, *Physical Origins of Entropy Production, Free Energy Dissipation, and Their Mathematical Representations*, Phys. Rev. E **81**, 051133 (2010).
[10] Massimiliano Esposito and Christian Van den Broeck, *Three Faces of the Second Law. I. Master Equation Formulation*, Phys. Rev. E **82**, 011143 (2010).
[11] Christian Van den Broeck and Massimiliano Esposito, *Three Faces of the Second Law. II. Fokker-Planck Formulation*, Phys. Rev. E **82**, 011144 (2010).
[12] Ken Sekimoto, *Langevin Equation and Thermodynamics*, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[13] Tânia Tomé and Mário J. de Oliveira, *Entropy Production in Irreversible Systems Described by a Fokker-Planck Equation*, Phys. Rev. E **82**, 021120 (2010).
[14] Ekaterina A. Korobkova, Thierry Emonet, Heungwon Park, and Philippe Cluzel, *Hidden Stochastic Nature of a Single Bacterial Motor*, Phys. Rev. Lett. **96**, 058105 (2006).
[15] Stefano Bo and Antonio Celani, *Multiple-Scale Stochastic Processes: Decimation, Averaging and Beyond*, Phys. Rep. **670**, 1 (2017).
[16] Massimiliano Esposito, *Stochastic Thermodynamics under Coarse Graining*, Phys. Rev. E **85**, 041125 (2012).
[17] Simone Pigolotti and Angelo Vulpiani, *Coarse Graining of Master Equations with Fast and Slow States*, J. Chem. Phys. **128**, 154114 (2008).
[18] Saar Rahav and Christopher Jarzynski, *Fluctuation Relations and Coarse-Graining*, J. Stat. Mech. (2007) P09012.
[19] Gianluca Teza and Attilio L. Stella, *Exact Coarse Graining Preserves Entropy Production out of Equilibrium*, Phys. Rev. Lett. **125**, 110601 (2020).
[20] Lucas Lacasa, Inés P. Mariño, Joaquin Miguez, Vincenzo Nicosia, Édgar Roldán, Ana Lisica, Stephan W. Grill, and Jesús Gómez-Gardeñes, *Multiplex Decomposition of Non-Markovian Dynamics and the Hidden Layer Reconstruction Problem*, Phys. Rev. X **8**, 031038 (2018).

[21] Ronald D. Vale, Thomas S. Reese, and Michael P. Sheetz, *Identification of a Novel Force-Generating Protein, Kinesin, Involved in Microtubule-Based Motility*, Cell **42**, 39 (1985).

[22] Jordanka Zlatanova and Kensal van Holde, *Single-Molecule Biology: What Is It and How Does It Work?*, Mol. Cell. **24**, 317 (2006).

[23] Sander Verbrugge, Lukas C. Kapitein, and Erwin J. G. Peterman, *Kinesin Moving through the Spotlight: Single-Motor Fluorescence Microscopy with Submillisecond Time Resolution*, Biophys. J. **92**, 2536 (2007).

[24] Vaishnavi Ananthanarayanan and Iva M. Tolić, *Single-Molecule Imaging of Cytoplasmic Dynein In Vivo*, in *Methods in Cell Biology* (Elsevier, New York, 2015), pp. 1–12.

[25] Stefan Niekamp, Nico Stuurman, Nan Zhang, and Ronald D. Vale, *Three-Color Single-Molecule Imaging Reveals Conformational Dynamics of Dynein Undergoing Motility*, Proc. Natl. Acad. Sci. U.S.A. **118** (2021).

[26] Rama Desai, Michael A. Geeves, and Neil M. Kad, *Using Fluorescent Myosin to Directly Visualize Cooperative Activation of Thin Filaments*, J. Biol. Chem. **290**, 1915 (2015).

[27] W. E. Moerner and David P. Fromm, *Methods of Single-Molecule Fluorescence Spectroscopy and Microscopy*, Rev. Sci. Instrum. **74**, 3597 (2003).

[28] Chirlmin Joo, Hamza Balci, Yuji Ishitsuka, Chittanon Buranachai, and Taekjip Ha, *Advances in Single-Molecule Fluorescence Methods for Molecular Biology*, Annu. Rev. Biochem. **77**, 51 (2008).

[29] Pierre Mangeol, Bram Prevo, and Erwin J. G. Peterman, *Kymographclear and Kymographdirect: Two Tools for the Automated Quantitative Analysis of Molecular and Cellular Dynamics Using Kymographs*, Mol. Biol. Cell **27**, 1948 (2016).

[30] Samara L. Reck-Peterson, Ahmet Yildiz, Andrew P. Carter, Arne Gennerich, Nan Zhang, and Ronald D. Vale, *Single-Molecule Analysis of Dynein Processivity and Stepping Behavior*, Cell **126**, 335 (2006).

[31] Elio A. Abbondanzieri, William J. Greenleaf, Joshua W. Shaevitz, Robert Landick, and Steven M. Block, *Direct Observation of Base-Pair Stepping by RNA Polymerase*, Nature (London) **438**, 460 (2005).

[32] Colin Echeverría Aitken, Alexey Petrov, and Joseph D. Puglisi, *Single Ribosome Dynamics and the Mechanism of Translation*, Annu. Rev. Biophys. **39**, 491 (2010).

[33] Jin-Der Wen, Laura Lancaster, Courtney Hodges, Ana-Carolina Zeri, Shige H. Yoshimura, Harry F. Noller, Carlos Bustamante, and Ignacio Tinoco, *Following Translation by Single Ribosomes One Codon at a Time*, Nature (London) **452**, 598 (2008).

[34] Annwesha Dutta, Gunter M. Schütz, and Debashish Chowdhury, *Stochastic Thermodynamics and Modes of Operation of a Ribosome: A Network Theoretic Perspective*, Phys. Rev. E **101**, 032402 (2020).

[35] Reinhard Lipowsky, Steffen Liepelt, and Angelo Valleriani, *Energy Conversion by Molecular Motors Coupled to Nucleotide Hydrolysis*, J. Stat. Phys. **135**, 951 (2009).

[36] Dominic J. Skinner and Jörn Dunkel, *Improved Bounds on Entropy Production in Living Systems*, Proc. Natl. Acad. Sci. U.S.A. **118** (2021).

[37] Shun Otsubo, Sreekanth K. Manikandan, Takahiro Sagawa, and Supriya Krishnamurthy, *Estimating Time-Dependent Entropy Production from Non-equilibrium Trajectories*, Communications in Physics **5** (2022).

[38] Matteo Polettini and Massimiliano Esposito, *Effective Fluctuation and Response Theory*, J. Stat. Phys. **176**, 94 (2019).

[39] Ignacio A. Martínez, Gili Bisker, Jordan M. Horowitz, and Juan M. R. Parrondo, *Inferring Broken Detailed Balance in the Absence of Observable Currents*, Nat. Commun. **10**, 3542 (2019).

[40] Naoto Shiraishi and Takahiro Sagawa, *Fluctuation Theorem for Partially Masked Nonequilibrium Dynamics*, Phys. Rev. E **91**, 012130 (2015).

[41] Notice that for any two states i and j, we have $\langle i|j\rangle = \delta_{i,j}$, with $\delta_{i,j}$ Kronecker's delta, whereas transitions $\langle\!\langle \ell||\ell \rangle\!\rangle \neq 1$.

[42] We denote by $D[P(x)||Q(x)] = \int_0^\infty \mathrm{d}x P(x) \ln[P(x)/Q(x)] \geq 0$ the Kullback-Leibler divergence between the probability distributions P and Q of the random variable x [43]. This information-theoretic measure can be generalized to distributions of multiple random variables and path probabilities of stochastic processes; see, e.g., Refs. [44–46] for applications in stochastic thermodynamics.

[43] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, in Wiley Series in Telecommunications and Signal Processing, 2nd ed. (Wiley-Interscience, New York, 2006).

[44] Édgar Roldán and Juan M. R. Parrondo, *Entropy Production and Kullback-Leibler Divergence between Stationary Trajectories of Discrete Systems*, Phys. Rev. E **85**, 031129 (2012).

[45] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, *Dissipation: The Phase-Space Perspective*, Phys. Rev. Lett. **98**, 080602 (2007).

[46] J. M. R. Parrondo, C. Van den Broeck, and R. Kawai, *Entropy Production and the Arrow of Time*, New J. Phys. **11**, 073008 (2009).

[47] Name suggested by the second referee.

[48] Sidney Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, England, 2001).

[49] Ken Sekimoto, *Derivation of the First Passage Time Distribution for Markovian Process on Discrete Network*, arXiv:2110.02216.

[50] Marco Baiesi, Christian Maes, and Bram Wynants, *Nonequilibrium Linear Response for Markov Dynamics, I: Jump Processes and Overdamped Diffusions*, J. Stat. Phys. **137**, 1094 (2009).

[51] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, *Dynamical First-Order Phase Transition in Kinetically Constrained Models of Glasses*, Phys. Rev. Lett. **98**, 195702 (2007).

[52] Gili Bisker, Matteo Polettini, Todd R. Gingrich, and Jordan M. Horowitz, *Hierarchical Bounds on Entropy Production Inferred from Partial Information*, J. Stat. Mech. (2017) 093210.

[53] Andre C. Barato and Udo Seifert, *Thermodynamic Uncertainty Relation for Biomolecular Processes*, Phys. Rev. Lett. **114**, 158101 (2015).
[54] Todd R. Gingrich, Jordan M. Horowitz, Nikolay Perunov, and Jeremy L. England, *Dissipation Bounds All Steady-State Current Fluctuations*, Phys. Rev. Lett. **116**, 120601 (2016).
[55] Jann van der Meer, Benjamin Ertel, and Udo Seifert, *Thermodynamic Inference in Partially Accessible Markov Networks: A Unifying Perspective from Transition-Based Waiting Time Distributions*, Phys. Rev. X **12**, 031025 (2022).
[56] Christian Maes, *Frenesy: Time-Symmetric Dynamical Activity in Nonequilibria*, Phys. Rep. **850**, 1 (2020).
[57] Édgar Roldán and Pierpaolo Vivo, *Exact Distributions of Currents and Frenesy for Markov Bridges*, Phys. Rev. E **100**, 042108 (2019).
[58] Christian Maes and Karel Netočný, *Nonequilibrium Corrections to Gradient Flow*, Chaos **29**, 073109 (2019).
[59] Yann R. Chemla, Jeffrey R. Moffitt, and Carlos Bustamante, *Exact Solutions for Kinetic Models of Macromolecular Dynamics*, J. Phys. Chem. B **112**, 6025 (2008).
[60] Juan P. Garrahan, *Simple Bounds on Fluctuations and Uncertainty Relations for First-Passage Times of Counting Observables*, Phys. Rev. E **95**, 032134 (2017).
[61] Pedro E. Harunari, Alberto Garilli, and Matteo Polettini, *The Beat of a Current*, arXiv:2205.05060.
[62] Matteo Polettini, *Fisher Information of Markovian Decay Modes*, Eur. Phys. J. B **87**, 215 (2014).
[63] Hong Qian and X. Sunney Xie, *Generalized Haldane Equation and Fluctuation Theorem in the Steady-State Cycle Kinetics of Single Enzymes*, Phys. Rev. E **74**, 010902 (R) (2006).
[64] Hao Ge, *Waiting Cycle Times and Generalized Haldane Equality in the Steady-State Cycle Kinetics of Single Enzymes*, J. Phys. Chem. B **112**, 61 (2008).
[65] Izaak Neri, Édgar Roldán, and Frank Jülicher, *Statistics of Infima and Stopping Times of Entropy Production and Applications to Active Molecular Processes*, Phys. Rev. X **7**, 011019 (2017).
[66] Matteo Polettini, *BEST Statistics of Markovian Fluxes: A Tale of Eulerian Tours and Fermionic Ghosts*, J. Phys. A **48**, 365005 (2015).
[67] Édgar Roldán and Juan M. R. Parrondo, *Estimating Dissipation from Single Stationary Trajectories*, Phys. Rev. Lett. **105**, 150607 (2010).
[68] J. Schnakenberg, *Network Theory of Microscopic and Macroscopic Behavior of Master Equation Systems*, Rev. Mod. Phys. **48**, 571 (1976).
[69] A. Gomez-Marin, J. M. R. Parrondo, and C. Van den Broeck, *Lower Bounds on Dissipation upon Coarse Graining*, Phys. Rev. E **78**, 011107 (2008).
[70] Fernando Perez-Cruz, *Kullback-Leibler Divergence Estimation of Continuous Distributions*, in *2008 IEEE International Symposium on Information Theory* (2008), pp. 1666–1670.
[71] Pedro E. Harunari and Ariel Yssou, *Kullback-Leibler Divergence Estimation Algorithm and Inter-transition Times Application*, https://github.com/pedroharunari/KLD_estimation (2022).
[72] Jos A. Morin, Francisco J. Cao, Jos M. Lzaro, J. Ricardo Arias-Gonzalez, Jos M. Valpuesta, Jos L. Carrascosa, Margarita Salas, and Borja Ibarra, *Mechano-Chemical Kinetics of DNA Replication: Identification of the Translocation Step of a Replicative DNA Polymerase*, Nucleic Acids Res. **43**, 3643 (2015).
[73] Xinghua Shi and Taekjip Ha, *Single-Molecule FRET: Technique and Applications to the Studies of Molecular Machines*, in Molecular Machines in Biology, edited by Joachim Frank (Cambridge University Press, Cambridge, England, 2011), pp. 4–19.
[74] Carlos J. Bustamante, Yann R. Chemla, Shixin Liu, and Michelle D. Wang, *Optical Tweezers in Single-Molecule Biophysics*, Nat. Rev. Meth. Primers **1** (2021).
[75] John T. Canty, Ruensern Tan, Emre Kusakci, Jonathan Fernandes, and Ahmet Yildiz, *Structure and Mechanics of Dynein Motors*, Annu. Rev. Biophys. **50**, 549 (2021).
[76] Jonathon Howard and R. L. Clark, *Mechanics of Motor Proteins and the Cytoskeleton*, Appl. Mech. Rev. **55**, B39 (2002).
[77] Andreja Šarlah and Andrej Vilfan, *The Winch Model Can Explain Both Coordinated and Uncoordinated Stepping of Cytoplasmic Dynein*, Biophys. J. **107**, 662 (2014).
[78] Wonseok Hwang and Changbong Hyeon, *Energetic Costs, Precision, and Transport Efficiency of Molecular Motors*, J. Phys. Chem. Lett. **9**, 513 (2018).
[79] Steffen Liepelt and Reinhard Lipowsky, *Kinesin's Network of Chemomechanical Motor Cycles*, Phys. Rev. Lett. **98**, 258102 (2007).
[80] Carlos Bustamante, Jan Liphardt, and Felix Ritort, *The Nonequilibrium Thermodynamics of Small Systems*, Phys. Today **58**, No. 07, 43 (2005).
[81] Nick J. Carter and R. A. Cross, *Mechanics of the Kinesin Step*, Nature (London) **435**, 308 (2005).
[82] Masayoshi Nishiyama, Hideo Higuchi, and Toshio Yanagida, *Chemomechanical Coupling of the Forward and Backward Steps of Single Kinesin Molecules*, Nat. Cell Biol. **4**, 790 (2002).
[83] Patrick Pietzonka, Andre C. Barato, and Udo Seifert, *Universal Bound on the Efficiency of Molecular Motors*, J. Stat. Mech. (2016) 124004.
[84] Andreas Dechant and Shin ichi Sasa, *Current Fluctuations and Transport Efficiency for General Langevin Systems*, J. Stat. Mech. (2018) 063209.
[85] Pierre Gaspard, *Template-Directed Copolymerization, Random Walks Along Disordered Tracks, and Fractals*, Phys. Rev. Lett. **117**, 238101 (2016).
[86] Sophia Rudorf and Reinhard Lipowsky, *Protein Synthesis in E. coli: Dependence of Codon-Specific Elongation on tRNA Concentration and Codon Usage*, PLoS One **10**, e0134994 (2015).
[87] Yariv Kafri, David K. Lubensky, and David R. Nelson, *Dynamics of Molecular Motors with Finite Processivity on Heterogeneous Tracks*, Phys. Rev. E **71**, 041906 (2005).
[88] Thomas Harms and Reinhard Lipowsky, *Driven Ratchets with Disordered Tracks*, Phys. Rev. Lett. **79**, 2895 (1997).
[89] Dominic J. Skinner and Jörn Dunkel, *Estimating Entropy Production from Waiting Time Distributions*, Phys. Rev. Lett. **127**, 198101 (2021).

[90] Yuhai Tu, *The Nonequilibrium Mechanism for Ultrasensitivity in a Biological Switch: Sensing by Maxwell's Demons*, Proc. Natl. Acad. Sci. U.S.A. **105**, 11737 (2008).

[91] Izaak Neri, Édgar Roldán, Simone Pigolotti, and Frank Jülicher, *Integral Fluctuation Relations for Entropy Production at Stopping Times*, J. Stat. Mech. (2019) 104006.

[92] David Hartich and Alja Godec, *Comment on "Inferring Broken Detailed Balance in the Absence of Observable Currents"*, arXiv:2112.08978.

[93] David Hartich and Alja ž. Godec, *Emergent Memory and Kinetic Hysteresis in Strongly Driven Networks*, Phys. Rev. X **11**, 041047 (2021).

[94] Jonathan Schubert, Andrea Schulze, Chrisostomos Prodromou, and Hannes Neuweiler, *Two-Colour Single-Molecule Photoinduced Electron Transfer Fluorescence Imaging Microscopy of Chaperone Dynamics*, Nat. Commun. **12**, 6964 (2021)..

[95] Janghyun Yoo, John M. Louis, Irina V. Gopich, and Hoi Sung Chung, *Three-Color Single-Molecule FRET and Fluorescence Lifetime Analysis of Fast Protein Folding*, J. Phys. Chem. B **122**, 11702 (2018).

[96] Kay Brandner and Udo Seifert, *Multi-terminal Thermoelectric Transport in a Magnetic Field: Bounds on Onsager Coefficients and Efficiency*, New J. Phys. **15**, 105003 (2013).

[97] Hadrien Vroylandt, Anthony Bonfils, and Gatien Verley, *Efficiency Fluctuations of Small Machines with Unknown Losses*, Phys. Rev. E **93**, 052123 (2016).

[98] Kristopher D. Rawls, Bonnie V. Dougherty, Edik M. Blais, Ethan Stancliffe, Glynis L. Kolling, Kalyan Vinnakota, Venkat R. Pannala, Anders Wallqvist, and Jason A. Papin, *A Simplified Metabolic Network Reconstruction to Promote Understanding and Development of Flux Balance Analysis Tools*, Computers in Biology and Medicine **105**, 64 (2019).

[99] Meera Sampath, Raja Sengupta, Stephane Lafortune, Kasim Sinnamohideen, and Demosthenis C. Teneketzis, *Failure Diagnosis Using Discrete-Event Models*, IEEE Trans. Control Syst. Technol. **4**, 105 (1996).

[100] Weilin Wang, Stéphane Lafortune, and Feng Lin, *An Algorithm for Calculating Indistinguishable States and Clusters in Finite-State Automata with Partially Observable Transitions*, Systems and Control Letters **56**, 656 (2007).

[101] William S. Lovejoy, *A Survey of Algorithmic Methods for Partially Observed Markov Decision Processes*, Ann. Oper. Res. **28**, 47 (1991).

[102] Anindya Bhadra, Edward L. Ionides, Karina Laneri, Mercedes Pascual, Menno Bouma, and Ramesh C. Dhiman, *Malaria in Northwest India: Data Analysis via Partially Observed Stochastic Differential Equation Models Driven by Lévy Noise*, J. Am. Stat. Assoc. **106**, 440 (2011).

[103] Viviana Serreli, Chin-Fa Lee, Euan R. Kay, and David A. Leigh, *A Molecular Information Ratchet*, Nature (London) **445**, 523 (2007).

[104] Klaara L. Viisanen, Samu Suomela, Simone Gasparinetti, Olli-Pentti Saira, Joachim Ankerhold, and Jukka P. Pekola, *Incomplete Measurement of Work in a Dissipative Two Level System*, New J. Phys. **17**, 055014 (2015).

[105] Massimo Borrelli, Jonne V. Koski, Sabrina Maniscalco, and Jukka P. Pekola, *Fluctuation Relations for Driven Coupled Classical Two-Level Systems with Incomplete Measurements*, Phys. Rev. E **91**, 012145 (2015).

[106] Philipp Strasberg, Gernot Schaller, Tobias Brandes, and Massimiliano Esposito, *Thermodynamics of a Physical Model Implementing a Maxwell Demon*, Phys. Rev. Lett. **110**, 040601 (2013).

[107] Terrell L. Hill, *Interrelations between Random Walks on Diagrams (Graphs) with and without Cycles*, Proc. Natl. Acad. Sci. U.S.A. **85**, 2879 (1988).

[108] Mamata Sahoo and Stefan Klumpp, *Backtracking Dynamics of RNA Polymerase: Pausing and Error Correction*, J. Phys. Condens. Matter **25**, 374104 (2013).

[109] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger, *Estimating Mutual Information*, Phys. Rev. E **69**, 066138 (2004).

[110] Juan A. Bonachela, Haye Hinrichsen, and Miguel A. Muñoz, *Entropy Estimates of Small Data Sets*, J. Phys. A **41**, 202001 (2008).