

This is the accepted manuscript made available via CHORUS. The article has been published as:

Conditional reversibility in nonequilibrium stochastic systems

Marcus V. S. Bonança and Christopher Jarzynski

Phys. Rev. E **93**, 022101 — Published 1 February 2016

DOI: [10.1103/PhysRevE.93.022101](https://doi.org/10.1103/PhysRevE.93.022101)

Conditional reversibility in nonequilibrium stochastic systems

Marcus V. S. Bonança*

Instituto de Física 'Gleb Wataghin', Universidade Estadual de Campinas, 13083-859, Campinas, São Paulo, Brazil

Christopher Jarzynski†

*Department of Chemistry and Biochemistry and Institute of Physical Sciences and Technology,
University of Maryland, College Park, Maryland 20742, USA*

(Dated: January 11, 2016)

For discrete-state stochastic systems obeying Markovian dynamics, we establish the counterpart of the *conditional reversibility theorem* obtained by Gallavotti for deterministic systems [Ann. de l'Institut Henri Poincaré (A) 70, 429 (1999)]. Our result states that stochastic trajectories conditioned on opposite values of entropy production are related by time reversal, in the long-time limit. In other words, the probability of observing a particular sequence of events, given a long trajectory with a specified entropy production rate σ , is the same as the probability of observing the time-reversed sequence of events, given a trajectory conditioned on the opposite entropy production, $-\sigma$, where both trajectories are sampled from the same underlying Markov process. To obtain our result, we use an equivalence between conditioned (“microcanonical”) and biased (“canonical”) ensembles of nonequilibrium trajectories. We provide an example to illustrate our findings.

PACS numbers: 05.70.Ln, 05.10.Gg, 02.50.Ga

* mbonanca@ifi.unicamp.br

† cjarzyns@umd.edu

I. INTRODUCTION

In equilibrium statistical mechanics, the equivalence of ensembles in the thermodynamic limit provides a useful tool for analyzing systems that are subject to sharp constraints. For example, for purposes of calculating averages in physical situations in which the total energy is fixed, the microcanonical ensemble can be replaced by the typically more convenient, fixed-temperature canonical ensemble, provided the temperature is chosen appropriately [1, 2]. Alternatively, if we wish to construct an ensemble that describes an equilibrium system conditioned on a specific value of an extensive observable, then we can use Boltzmann-like weights to bias the distribution toward that value; in the thermodynamic limit the conditioned and biased distributions become equivalent. The mathematical tools for constructing such ensembles have been analyzed rigorously within the theory of large deviations [3].

In the last decade these tools have been applied extensively to probability distributions $\mathcal{P}[S]$ on the space of *paths* or trajectories of nonequilibrium systems [4–10]. This approach has been used to study dynamical phase transitions in kinematically constrained models [11–13], glass transitions [14–16], quantum systems [17–19], and efficiency fluctuations in stochastic heat engines [20–22]. By analogy with the equilibrium case, nonequilibrium path ensembles can be constructed by introducing exponential, Boltzmann-like weights $e^{-\lambda \mathcal{A}[S]}$ to modify a given probability distribution. Here the quantity inside the exponent is a *time*-extensive functional of the trajectory, $\mathcal{A}[S]$, multiplied by a biasing parameter $-\lambda$. The formal analogy with the equilibrium case suggests that such biased ensembles may be equivalent, in the appropriate limit, to nonequilibrium ensembles conditioned on specified values of the path observable \mathcal{A} . This problem has been addressed explicitly by Jack and Sollich [8] for discrete-state systems, and by Chetrite and Touchette [9, 10] for a broad class of stochastic models including diffusive processes. Using the theory of large deviations, these authors have established the equivalence between conditioned (“microcanonical”) and biased (“canonical”) path space distributions, in the long-time limit, given certain conditions related to the fluctuations of the observable appearing in the exponential weight. The results contained in Refs. [8–10] are related to earlier results by Evans [4, 5, 7], obtained within the framework of maximum-entropy inference, as well as to the generalized Onsager-Machlup theory developed by Maes and Netočný [6].

Here we use this equivalence to study nonequilibrium systems that are conditioned on values of entropy production. For a discrete-state system whose time-averaged rate of entropy production takes a positive value $\bar{\sigma}$ in the infinite-time limit, Eq. (9), we consider the statistical fluctuations in the entropy production rate over *finite* intervals of duration \mathcal{T} . Heuristically, it is useful to imagine “chopping” an infinite trajectory into segments of duration \mathcal{T} , and then segregating these according to the time-averaged entropy production rate during each segment. Those segments for which the entropy production rate takes on a value σ comprise an ensemble that is conditioned on that value. In the long- \mathcal{T} limit, we find that ensembles conditioned on *opposite* values of entropy production rates are related by time-reversal. In effect, if we compare two long trajectory segments, conditioned on entropy production rates $\pm\sigma$, then (statistically) one of them will look like a mirror image of the other, in time. This result is the stochastic counterpart of the *conditional reversibility theorem* derived by Gallavotti [23] in the context of deterministic dynamics.

In our presentation we will not aim at full mathematical rigor, nor will we assume that the reader is deeply familiar with large deviation theory. We will show that our central result, which is the conditional reversibility described above, follows from relatively straightforward manipulations. In order to keep the presentation self-contained, in Sec. II we will derive results obtained previously in Refs. [8–10].

The outline of the paper is as follows. In Sec. II we define a stationary, discrete-state Markov process that violates detailed balance, Eq. (1); and following [8–10] we construct the rate matrix for the biased ensemble, Eq. (37). In Sec. III we establish that ensembles biased toward opposite values of entropy production rates $\pm\sigma$ are described by rate matrices that are the dual of one another, Eqs. (46), (48), and we use this result to formulate a conditional reversibility theorem for stochastic dynamics, Eq. (53). In Sec. IV we present an illustrative example and in Sec. V we finish with concluding remarks.

II. CONDITIONED AND BIASED ENSEMBLES AND THEIR DYNAMICS

We consider a model with K discrete states labelled by $i \in \Lambda \equiv \{1, \dots, K\}$. The probability $p_i(\tau)$ to find the system in state i at time τ evolves according to the master equation

$$\frac{dp_i}{d\tau} = \sum_{j \neq i}^K (R_{ij} p_j - R_{ji} p_i) = \sum_{j=1}^K R_{ij} p_j \quad (1)$$

where the transition rates R_{ij} from state j to state i are time-independent, and $\sum_{j \neq i} R_{ji} = -R_{ii}$ is the escape rate from state i . The formal solution of (1) reads

$$\mathbf{p}(\tau) = e^{(\tau - \tau_0) R} \mathbf{p}(\tau_0), \quad (2)$$

where $\mathbf{p} \equiv (p_1, \dots, p_K)^T$ and $\mathbf{p}(\tau_0)$ is an initial probability distribution.

We assume that the K states form a connected network: from any state i the system can evolve to any other state j by a finite sequence of transitions. We also assume that if $R_{ij} \neq 0$, then $R_{ji} \neq 0$. These assumptions imply a unique stationary distribution, which we will denote by $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)^T$:

$$\lim_{\tau \rightarrow \infty} \mathbf{p}(\tau) = \boldsymbol{\pi}. \quad (3)$$

This distribution is characterized by stationary currents,

$$J_{ij} = R_{ij}\pi_j - R_{ji}\pi_i \quad (4)$$

representing the flow of probability from j to i in the stationary state.

We denote by $\mathcal{P}[\mathcal{S}]$ the probability of observing a trajectory \mathcal{S} over an interval of total duration \mathcal{T} , beginning at time $\tau_0 = -\mathcal{T}/2$ and ending at time $\mathcal{T}/2$. Discretizing time in intervals of duration $\delta\tau$, we represent this trajectory as a sequence of states $\mathcal{S} = (s_{-N}, s_{-N+1}, \dots, s_{N-1}, s_N) \in \Lambda^{2N+1}$, where s_n is the state of the system at time $n\delta\tau$ and $\delta\tau = \mathcal{T}/2N$. We will take $\delta\tau$ to be infinitesimal, hence $\delta\tau R_{ij} \ll 1$ for all i, j , and we will ignore terms of order $\delta\tau^2$. The probability $\mathcal{P}[\mathcal{S}]$ can be written as

$$\mathcal{P}[\mathcal{S}] = \left(\prod_{k=-N}^{N-1} U_{s_{k+1}s_k} \right) p_{s_{-N}}(\tau_0), \quad (5)$$

where the transition probability U_{ij} is given by

$$U_{ij} = (e^{\delta\tau R})_{ij} \simeq \delta_{ij} + \delta\tau R_{ij}. \quad (6)$$

The distribution $\mathcal{P}[\mathcal{S}]$ defines a statistical ensemble of discretized trajectories of duration \mathcal{T} .

The time-averaged entropy production rate along a trajectory \mathcal{S} is given by

$$\sigma[\mathcal{S}] = \frac{1}{\mathcal{T}} \sum_{k=-N}^{N-1} \ln \left(\frac{R_{s_{k+1}s_k}}{R_{s_k s_{k+1}}} \right), \quad (7)$$

where $\ln(R_{ij}/R_{ji})$ is the entropy production associated with a transition from j to i [24–27]. The probability distribution of the time-averaged entropy production rate is then

$$P(\sigma) = \sum_{[\mathcal{S}]} \mathcal{P}[\mathcal{S}] \delta(\sigma - \sigma[\mathcal{S}]), \quad (8)$$

where $\sum_{[\mathcal{S}]} \equiv \sum_{s_{-N}, \dots, s_N}$. Finally, we define a *conditioned* ensemble of trajectories, described by a probability distribution $\mathcal{P}_c[\mathcal{S}|\sigma]$, which is the subensemble of $\mathcal{P}[\mathcal{S}]$ containing only those trajectories with a time-averaged entropy production rate equal to σ . Our aim is to study the dynamics of the system within this conditioned ensemble.

Note that the distributions $\mathcal{P}[\mathcal{S}]$, $P(\sigma)$ and $\mathcal{P}_c[\mathcal{S}|\sigma]$ – as well as $\mathcal{P}_\lambda[\mathcal{S}]$ and $P_\lambda(\sigma)$, defined below – all depend on the duration \mathcal{T} , but this dependence is notationally suppressed. We are interested in the long-time limit, $\mathcal{T} \rightarrow \infty$ (with $\delta\tau$ fixed, hence $N \rightarrow \infty$). In this limit, the distribution $P(\sigma)$ becomes ever more sharply peaked around a value $\bar{\sigma}$, which is the infinite-time average entropy production rate in the stationary state:

$$\bar{\sigma} = \sum_{i>j} J_{ij} \ln \left(\frac{R_{ij}}{R_{ji}} \right). \quad (9)$$

Now consider a *biased* ensemble, described by the probability distribution

$$\mathcal{P}_\lambda[\mathcal{S}] = \frac{1}{\mathcal{N}} \mathcal{P}[\mathcal{S}] e^{-\lambda \mathcal{T} \sigma[\mathcal{S}]} \quad (10)$$

where λ is a real parameter and

$$\mathcal{N} = \mathcal{N}(\lambda) = \sum_{[\mathcal{S}]} \mathcal{P}[\mathcal{S}] e^{-\lambda \mathcal{T} \sigma[\mathcal{S}]} \quad (11)$$

is a normalization factor. The distribution $\mathcal{P}_\lambda[\mathcal{S}]$ defines a modified distribution of entropy production rates

$$P_\lambda(\sigma) = \sum_{[\mathcal{S}]} \mathcal{P}_\lambda[\mathcal{S}] \delta(\sigma - \sigma[\mathcal{S}]). \quad (12)$$

When $\lambda > 0$, the factor $e^{-\lambda \mathcal{T} \sigma[\mathcal{S}]}$ favors trajectories with low values of σ , hence the distribution $P_\lambda(\sigma)$ is shifted to the left of $P(\sigma)$; the opposite comments apply when $\lambda < 0$.

As shown in Refs. [8–10], in the long-time limit the statistics of the biased ensemble become equivalent to those of a stationary Markov process, described by a rate matrix $\tilde{R}(\lambda)$, Eq. (37). Moreover, in this limit the distribution $P_\lambda(\sigma)$ becomes sharply peaked around a value $\bar{\sigma}_\lambda$. This value decreases monotonically with λ , and can be written as

$$\bar{\sigma}_\lambda = \sum_{i>j} \tilde{J}_{ij}(\lambda) \ln \left(\frac{R_{ij}}{R_{ji}} \right), \quad (13)$$

where $\tilde{J}_{ij}(\lambda)$ is the net flow of probability from j to i in the biased λ -ensemble, in the long-time limit. Note that in Eqs. (12) and (13), the entropy production rate is defined with respect to the original, unbiased transition rates R_{ij} .

We can view the transformation $\mathcal{P}[\mathcal{S}] \rightarrow \mathcal{P}[\mathcal{S}] e^{-\lambda \mathcal{T} \sigma[\mathcal{S}]}$ as a method for constructing ensembles of trajectories that are effectively conditioned on particular entropy production rates. By varying λ , we “tune in” to trajectories with values of σ near a desired value $\bar{\sigma}_\lambda$; the larger the value of \mathcal{T} , the narrower the distribution of values of σ around $\bar{\sigma}_\lambda$. Under conditions discussed by Jack and Sollich [8] and by Chetrite and Touchette [9, 10], which are fulfilled by the model we study, in the long-time limit the ensemble $\mathcal{P}_\lambda[\mathcal{S}]$ becomes equivalent to an ensemble in which the entropy production rate is constrained to the value $\bar{\sigma}_\lambda$. We represent this equivalence using the notation

$$\mathcal{P}_\lambda[\mathcal{S}] \sim \mathcal{P}_c[\mathcal{S}|\bar{\sigma}_\lambda], \quad (14)$$

where the limit $\mathcal{T} \rightarrow \infty$ is implied.

As in the equilibrium case, the equivalence of nonequilibrium ensembles expressed by Eq. (14) helps us to avoid the difficulties imposed by sharp constraints. For nonequilibrium path-space ensembles, the long-time limit $\mathcal{T} \rightarrow \infty$ is analogous to the thermodynamic limit, and we can select a value of λ that gives a desired entropy production rate $\bar{\sigma}_\lambda$: the long-time dynamics of the system in the λ -ensemble $\mathcal{P}_\lambda[\mathcal{S}]$ are equivalent to its dynamics in the corresponding fixed- σ ensemble, $\mathcal{P}_c[\mathcal{S}|\bar{\sigma}_\lambda]$. We will exploit this equivalence in order to explore the behavior of the system when its long-time-averaged entropy production rate is conditioned on a particular value.

In the remainder of this section we first establish that the λ -ensemble $\mathcal{P}_\lambda[\mathcal{S}]$ describes a stationary process whose dynamics are Markovian in the long-time limit, Eq. (31), and we obtain the rate matrix that generates this process, Eq. (37).

We begin with the expression

$$\mathcal{P}_\lambda[\mathcal{S}] = \mathcal{N}^{-1} \left(\prod_{k=-N}^{N-1} U_{s_{k+1}s_k} \right) e^{-\lambda \mathcal{T} \sigma[\mathcal{S}]} p_{s_{-N}}(\tau_0) = \mathcal{N}^{-1} \left[\prod_{k=-N}^{N-1} U_{s_{k+1}s_k} \left(\frac{R_{s_k s_{k+1}}}{R_{s_{k+1} s_k}} \right)^\lambda \right] p_{s_{-N}}(\tau_0), \quad (15)$$

which follows from Eqs. (5) and (10), along with

$$e^{-\lambda \mathcal{T} \sigma[\mathcal{S}]} = \prod_{k=-N}^{N-1} \left(\frac{R_{s_k s_{k+1}}}{R_{s_{k+1} s_k}} \right)^\lambda. \quad (16)$$

Defining a matrix $Q(\lambda)$ with elements $Q_{ij} \equiv U_{ij} (R_{ji}/R_{ij})^\lambda$, we rewrite Eq. (15) as

$$\mathcal{P}_\lambda[\mathcal{S}] = \mathcal{N}^{-1} \left[\prod_{k=-N}^{N-1} Q_{s_{k+1}s_k} \right] p_{s_{-N}}(\tau_0). \quad (17)$$

Moreover we can use Eq. (6) to obtain

$$Q_{ij} \simeq \begin{cases} 1 + R_{ii} \delta t, & i = j \\ R_{ij}^{1-\lambda} R_{ji}^\lambda \delta t, & i \neq j \end{cases}. \quad (18)$$

Next, we introduce the convenient bra-ket notation to denote the right and left eigenvectors of Q :

$$Q|\mathbf{u}_k\rangle = \alpha_k|\mathbf{u}_k\rangle, \quad , \quad \langle\mathbf{v}_k|Q = \alpha_k\langle\mathbf{v}_k|, \quad (19)$$

where α_k are the corresponding eigenvalues. For the sake of simplicity we assume Q to be diagonalizable, so that the following relations apply after suitable normalization of the eigenvectors:

$$\langle\mathbf{v}_k|\mathbf{u}_j\rangle = \delta_{kj}, \quad (20a)$$

$$\sum_k |\mathbf{u}_k\rangle\langle\mathbf{v}_k| = \mathbb{1}. \quad (20b)$$

In the appendix we extend our derivation to the case when Q is *defective*, that is non-diagonalizable [28].

Because the matrix Q is non-negative (by construction) and irreducible (since the K states form a connected network), the conditions of the Perron-Frobenius theorem are satisfied [29]. This theorem tells us that Q has a real eigenvalue μ whose value is greater than the modulus of any other eigenvalue:

$$\mu \equiv \alpha_1 > |\alpha_2| \geq |\alpha_3| \geq \dots \quad (21)$$

Without loss of generality we have arranged the α_k 's in descending order of their moduli. We denote by

$$|\psi\rangle \equiv |\mathbf{u}_1\rangle \quad \text{and} \quad \langle\omega| \equiv \langle\mathbf{v}_1| \quad (22)$$

the right and left eigenvectors associated with the eigenvalue μ . The Perron-Frobenius theorem further guarantees that the elements of these eigenvectors are strictly positive. Note that the eigenvectors and eigenvalues of $Q(\lambda)$ depend on the value of λ . In particular, when $\lambda = 0$ we have $Q = U$, $\mu = 1$, $|\psi\rangle = |\pi\rangle$ and $\langle\omega| = \langle\mathbf{1}|$, where $\langle\mathbf{1}| \equiv (1, 1, \dots, 1)$.

From Eq. (17) we have

$$\mathcal{N} = \sum_{[S]} \left[\prod_{k=-N}^{N-1} Q_{s_{k+1}s_k} \right] p_{s_{-N}}(\tau_0) \quad (23)$$

$$= \sum_{[S]} Q_{s_N s_{N-1}} Q_{s_{N-1} s_{N-2}} \cdots Q_{s_{-N+1} s_{-N}} p_{s_{-N}}(\tau_0) \quad (24)$$

$$= \langle\mathbf{1}|Q^{2N}|\mathbf{p}(\tau_0)\rangle \quad (25)$$

Introducing the notation $\mathbf{p}_0 \equiv \mathbf{p}(\tau_0)$ and using Eq. (20), we obtain

$$\mathcal{N} = \sum_k \langle\mathbf{1}|Q^{2N}|\mathbf{u}_k\rangle\langle\mathbf{v}_k|\mathbf{p}_0\rangle = \sum_k \alpha_k^{2N} \langle\mathbf{1}|\mathbf{u}_k\rangle\langle\mathbf{v}_k|\mathbf{p}_0\rangle \rightarrow \mu^{2N} \langle\mathbf{1}|\psi\rangle\langle\omega|\mathbf{p}_0\rangle \quad , \quad (26)$$

where \rightarrow denotes the long-time limit ($\mathcal{T}, N \rightarrow \infty$), and we have invoked Eq. (21) when taking this limit.

With these elements in place, let us consider the probability, in the λ -ensemble, that the system proceeds through a sequence of states $r_0, r_1, \dots, r_\kappa$ at times $0, \delta\tau, \dots, \kappa\delta\tau$. (Without loss of generality, we have taken $j = 0$ as the initial time step.) This probability is given by

$$p_\lambda(r_\kappa, \kappa; r_{\kappa-1}, \kappa-1; \dots; r_0, 0) = \mathcal{N}^{-1} \sum_{[S]} \left[\prod_{k=-N}^{N-1} Q_{s_{k+1}s_k} \right] p_{s_{-N}}(\tau_0) \delta_{s_0 r_0} \delta_{s_1 r_1} \cdots \delta_{s_\kappa r_\kappa}. \quad (27)$$

Here we have inserted appropriate Kronecker delta functions $\delta_{s_j r_j}$ into Eq. (17) and taken a sum over trajectories. Let us now introduce the notation $|\mathbf{e}_s\rangle = (\delta_{1s}, \delta_{2s}, \dots, \delta_{Ks})^T$ to denote a unit vector. Proceeding as in the previous paragraph we get

$$\begin{aligned} p_\lambda(r_\kappa, \kappa; \dots; r_0, 0) &= \mathcal{N}^{-1} \langle\mathbf{1}|Q^{N-\kappa}|\mathbf{e}_{r_\kappa}\rangle Q_{r_\kappa r_{\kappa-1}} \cdots Q_{r_1 r_0} \langle\mathbf{e}_{r_0}|Q^N|\mathbf{p}_0\rangle \\ &= \mathcal{N}^{-1} \sum_{l,m} \langle\mathbf{1}|Q^{N-\kappa}|\mathbf{u}_l\rangle\langle\mathbf{v}_l|\mathbf{e}_{r_\kappa}\rangle Q_{r_\kappa r_{\kappa-1}} \cdots Q_{r_1 r_0} \langle\mathbf{e}_{r_0}|Q^N|\mathbf{u}_m\rangle\langle\mathbf{v}_m|\mathbf{p}_0\rangle \\ &\rightarrow \mathcal{N}^{-1} \mu^{2N-\kappa} \langle\mathbf{1}|\psi\rangle\omega_{r_\kappa} \left(\prod_{j=0}^{\kappa-1} Q_{r_{j+1}r_j} \right) \psi_{r_0} \langle\omega|\mathbf{p}_0\rangle \\ &= \mu^{-\kappa} \omega_{r_\kappa} \left(\prod_{j=0}^{\kappa-1} Q_{r_{j+1}r_j} \right) \psi_{r_0} = \left(\prod_{j=0}^{\kappa-1} \frac{\omega_{r_{j+1}} Q_{r_{j+1}r_j} \omega_{r_j}^{-1}}{\mu} \right) \omega_{r_0} \psi_{r_0} \end{aligned} \quad (28)$$

For the special cases $\kappa = 0$ and $\kappa = 1$ we get

$$p_\lambda(r_0, 0) = \omega_{r_0} \psi_{r_0} \quad , \quad p_\lambda(r_1, 1; r_0, 0) = \frac{\omega_{r_1} Q_{r_1 r_0} \omega_{r_0}^{-1}}{\mu} \omega_{r_0} \psi_{r_0} \quad , \quad (29)$$

implying that the conditional probability to be found in state r_1 at $t = \delta\tau$, given state r_0 at $t = 0$, is

$$p_\lambda(r_1, 1 | r_0, 0) = \frac{\omega_{r_1} Q_{r_1 r_0} \omega_{r_0}^{-1}}{\mu}. \quad (30)$$

An identical result holds for $p_\lambda(r_{j+1}, j+1 | r_j, j)$ for any j (since the choice of $j = 0$ as the initial time step in Eq. (28) was arbitrary), allowing us to rewrite Eq. (28) as follows:

$$p_\lambda(r_\kappa, \kappa; \dots; r_0, 0) = \left[\prod_{j=0}^{\kappa-1} p_\lambda(r_{j+1}, j+1 | r_j, j) \right] p_\lambda(r_0, 0), \quad (31)$$

from which we conclude that the trajectory segment $r_0, r_1, \dots, r_\kappa$, sampled from the λ -ensemble, is described statistically by a Markov process, in the limit $\mathcal{T}, N \rightarrow \infty$.

Following Refs. [8–10], let us now construct the rate matrix that governs the λ -biased dynamics. As discussed in Refs. [9, 10], this rate matrix is related to the original, unbiased rate matrix by Doob's transform [30].

We have used the notation R and U to denote the rate matrix and the single-time-step transition matrix for the unbiased dynamics. In what follows we will use $\tilde{R}(\lambda)$ and $\tilde{U}(\lambda)$ to denote the rate and transition matrices for the biased dynamics, thus $\tilde{R}(0) = R$ and $\tilde{U}(0) = U$. The elements of $\tilde{U}(\lambda)$ are given by Eq. (30):

$$\tilde{U} = \frac{1}{\mu} \Omega Q \Omega^{-1} \quad (32)$$

where $\Omega \equiv \text{diag}(\langle \omega |)$ is a diagonal matrix whose elements are the components of the left eigenvector $\langle \omega |$ of Q . Using $\tilde{U} = e^{\tilde{R}\delta\tau} \simeq \mathbb{1} + \tilde{R}\delta\tau$ (where $\mathbb{1}$ is the identity matrix) we obtain

$$\tilde{R}\delta\tau = \frac{1}{\mu} \Omega Q \Omega^{-1} - \mathbb{1} \quad (33)$$

To rewrite this expression in a more convenient form, we first use Eq. (18) to write

$$Q = \mathbb{1} + M \delta\tau, \quad (34)$$

where

$$M_{ij} = \begin{cases} R_{ii}, & i = j \\ R_{ij}^{1-\lambda} R_{ji}^\lambda, & i \neq j. \end{cases} \quad (35)$$

By Eq. (34), Q and M have the same eigenvectors but different eigenvalues, and in particular

$$\mu = 1 + \eta \delta\tau \quad , \quad \langle \omega | M = \eta \langle \omega | \quad , \quad M | \psi \rangle = \eta | \psi \rangle. \quad (36)$$

Keeping terms only up to first order in $\delta\tau$, Eq. (33) gives us

$$\tilde{R}(\lambda) = \Omega M \Omega^{-1} - \eta \mathbb{1}, \quad (37)$$

where the quantities appearing on the right side of the equation depend on λ , but not on $\delta\tau$. This rate matrix, $\tilde{R}(\lambda)$, generates trajectories with the same statistics as those of the λ -ensemble, in the long-time limit.

Let us now solve for the null right eigenvector of the rate matrix $\tilde{R}(\lambda)$. We will use the notation $|\tilde{\pi}(\lambda)\rangle$ to denote this eigenvector, whose components $\tilde{\pi}_s = \langle \mathbf{e}_s | \tilde{\pi} \rangle$ represent the stationary probability distribution in the biased λ -ensemble. We have

$$|\mathbf{0}\rangle = \tilde{R}|\tilde{\pi}\rangle = \Omega M \Omega^{-1} |\tilde{\pi}\rangle - \eta |\tilde{\pi}\rangle \quad (38)$$

hence

$$M \left(\Omega^{-1} |\tilde{\pi}\rangle \right) = \eta \left(\Omega^{-1} |\tilde{\pi}\rangle \right), \quad (39)$$

which implies by Eq. (36) that $\Omega^{-1}|\tilde{\pi}\rangle \propto |\psi\rangle$. We thus have $|\tilde{\pi}\rangle = c\Omega|\psi\rangle$, or $\tilde{\pi}_s = \langle \mathbf{e}_s | \tilde{\pi} \rangle = c\omega_s\psi_s$. The constant c is set by the normalization condition $\sum_s \tilde{\pi}_s = 1$, which combines with $\langle \mathbf{v}_1 | \mathbf{u}_1 \rangle = 1$ to give us $c = 1$, hence

$$|\tilde{\pi}\rangle = \Omega|\psi\rangle \quad , \quad \tilde{\pi}_s = \omega_s\psi_s. \quad (40)$$

This is consistent with Eq. (29).

To end this section, we establish symmetry identities, Eqs. (41) and (42), that will allow us to investigate the dependence of the biased dynamics on the parameter λ . From Eq. (35) we have $M(1-\lambda) = M(\lambda)^T$, which implies

$$\eta(\lambda) = \eta(1-\lambda) \quad , \quad |\psi(\lambda)\rangle = |\omega(1-\lambda)\rangle. \quad (41)$$

These results combine with the transpose of Eq. (37) to give us

$$\Omega(\lambda)\tilde{R}(\lambda)^T\Omega(\lambda)^{-1} = M(1-\lambda) - \eta(1-\lambda)\mathbb{1} \quad (42)$$

The symmetry property embodied by Eqs. (41) and (42) was previously obtained by Lebowitz and Spohn [24], in their proof of the fluctuation theorem for general Markov processes. A similar symmetry is implicit in Kurchan's earlier derivation of the fluctuation theorem for Langevin processes [31].

III. DUAL DYNAMICS AND CONDITIONAL REVERSIBILITY

We are now in a position to study the *dual* of the rate matrix \tilde{R} , defined by [32, 33]

$$\tilde{R}^{\text{dual}} = \tilde{\Pi}\tilde{R}^T\tilde{\Pi}^{-1} \quad (43)$$

where $\tilde{\Pi} = \text{diag}(|\tilde{\pi}\rangle)$. From this definition it follows that $\langle \mathbf{1} | \tilde{R}^{\text{dual}} = \langle \mathbf{0} |$ and $\tilde{R}^{\text{dual}}|\tilde{\pi}\rangle = |\mathbf{0}\rangle$, hence \tilde{R}^{dual} is a rate matrix whose stationary distribution is the same as that of \tilde{R} . Moreover, in the stationary state the flow of probability under the dual dynamics is exactly opposite to the probability flow under the original dynamics:

$$\tilde{J}_{ij}^{\text{dual}} = \tilde{R}_{ij}^{\text{dual}}\tilde{\pi}_j - \tilde{R}_{ji}^{\text{dual}}\tilde{\pi}_i = \left(\tilde{\pi}_i\tilde{R}_{ji}\tilde{\pi}_j^{-1}\right)\tilde{\pi}_j - \left(\tilde{\pi}_j\tilde{R}_{ij}\tilde{\pi}_i^{-1}\right)\tilde{\pi}_i = -\tilde{J}_{ij}. \quad (44)$$

Combining Eq. (43) with Eqs. (37), (40), (41) and (42), we obtain

$$\begin{aligned} \left[\tilde{R}^{\text{dual}}(\lambda)\right]_{ij} &= \tilde{\pi}_i(\lambda) \left[\tilde{R}(\lambda)^T\right]_{ij} \tilde{\pi}_j(\lambda)^{-1} \\ &= \psi_i(\lambda) \left[\Omega(\lambda)\tilde{R}(\lambda)^T\Omega(\lambda)^{-1}\right]_{ij} \psi_j(\lambda)^{-1} \\ &= \psi_i(\lambda) [M(1-\lambda) - \eta(1-\lambda)\mathbb{1}]_{ij} \psi_j(\lambda)^{-1} \\ &= \omega_i(1-\lambda) [M(1-\lambda)]_{ij} \omega_j(1-\lambda)^{-1} - \eta(1-\lambda)\delta_{ij} = \left[\tilde{R}(1-\lambda)\right]_{ij} \end{aligned} \quad (45)$$

That is,

$$\tilde{R}^{\text{dual}}(\lambda) = \tilde{R}(1-\lambda). \quad (46)$$

Together with Eq. (44), this result gives us

$$\tilde{J}_{ij}(1-\lambda) = -\tilde{J}_{ij}(\lambda). \quad (47)$$

Thus the stationary currents in the λ -ensemble are the opposite of those in the $(1-\lambda)$ -ensemble, which further implies (see Eq. (13)) that the corresponding average entropy production rates are opposite:

$$\bar{\sigma}_{1-\lambda} = -\bar{\sigma}_\lambda. \quad (48)$$

Therefore, by the equivalence of ensembles expressed by Eq. (14), in the long-time limit the biased ensembles $\mathcal{P}_\lambda[\mathcal{S}]$ and $\mathcal{P}_{1-\lambda}[\mathcal{S}]$ become equivalent to ensembles *conditioned on opposite entropy production rates*, $\mathcal{P}_c[\mathcal{S}|\bar{\sigma}_\lambda]$ and $\mathcal{P}_c[\mathcal{S}|- \bar{\sigma}_\lambda]$.

In Sec. II we discussed the equivalence between *biased* and *conditioned* ensembles of trajectories; see Eq. (14). Combined with Eq. (46), this equivalence implies that ensembles conditioned on opposite values of entropy production, $\mathcal{P}_c[\mathcal{S}|\sigma]$ and $\mathcal{P}_c[\mathcal{S}|- \sigma]$, are characterized by the same stationary distribution, but opposite currents. This suggests that these ensembles may be related by time-reversal, not only with respect to time-averaged currents and entropy

production, but also at the level of individual trajectories. We now explore this idea in detail, to arrive at our central result, Eq. (53).

Let us consider an interval of time of finite duration $\kappa \delta\tau$, and let $\mathbf{r} = (r_0 \rightarrow r_1 \rightarrow \dots \rightarrow r_\kappa)$ denote a trajectory segment evolving during this interval, in which the system is found in state r_0 at time step 0, in state r_1 at time step 1, and so forth up to time step κ . The notation $\mathbf{r}^* = (r_\kappa \rightarrow \dots \rightarrow r_1 \rightarrow r_0)$ will denote the time-reversed segment, in which the same states are visited in reverse order. If we sample a trajectory \mathcal{S} from the biased λ -ensemble, and we examine the states visited by this trajectory from $\tau = 0$ to $\tau = \kappa \delta\tau$, then the probability to observe the trajectory segment \mathbf{r} is given by

$$P_\lambda(\mathbf{r}) = \tilde{\pi}_{r_0}(\lambda) \tilde{U}_{r_1 r_0}(\lambda) \cdots \tilde{U}_{r_\kappa r_{\kappa-1}}(\lambda) \rightarrow \left[\frac{1}{\mu(\lambda)} \right]^\kappa \omega_{r_\kappa}(\lambda) \left[\prod_{j=0}^{\kappa-1} Q_{r_{j+1} r_j}(\lambda) \right] \psi_{r_0}(\lambda), \quad (49)$$

using Eqs. (32) and (40). Similarly,

$$P_{1-\lambda}(\mathbf{r}^*) \rightarrow \left[\frac{1}{\mu(1-\lambda)} \right]^\kappa \omega_{r_0}(1-\lambda) \left[\prod_{j=0}^{\kappa-1} Q_{r_j r_{j+1}}(1-\lambda) \right] \psi_{r_\kappa}(1-\lambda). \quad (50)$$

Let us now compare these probabilities.

From Eq. (18) we have $Q(1-\lambda) = Q(\lambda)^T$, which implies [24]

$$\mu(1-\lambda) = \mu(\lambda) \quad , \quad \omega_s(1-\lambda) = \psi_s(\lambda) \quad (51)$$

(compare with Eq. (41)). Combining these results with Eqs. (49) and (50), we obtain

$$P_\lambda(\mathbf{r}) = P_{1-\lambda}(\mathbf{r}^*). \quad (52)$$

By the equivalence of ensembles, this implies the following conclusion. The probability to observe a given trajectory segment $\mathbf{r} = (r_0 \rightarrow \dots \rightarrow r_\kappa)$ when conditioning on a particular value of time-averaged entropy production rate σ , is the same as the probability to observe the time-reversed segment $\mathbf{r}^* = (r_\kappa \rightarrow \dots \rightarrow r_0)$ when conditioning on the opposite value of time-averaged entropy production rate, $-\sigma$. Using obvious notation:

$$P_c(\mathbf{r}|\sigma) = P_c(\mathbf{r}^*|-\sigma) \quad (53)$$

It is in this sense that the two conditioned ensembles are related by time-reversal.

Equation (52) is related to expressions that already exist in the literature. For instance, Eq. (37) of Ref. [25] and Eq. (2.21) of Ref. [24] can be written in our notation as (see Sec. II)

$$\mathcal{P}[\mathcal{S}] e^{-\lambda \mathcal{T} \sigma[\mathcal{S}]} = \frac{p_{s_{-N}}(\tau_0)}{p_{s_N}(\tau_f)} \mathcal{P}[\mathcal{S}^*] e^{-(1-\lambda) \mathcal{T} \sigma[\mathcal{S}^*]}, \quad (54)$$

where $\tau_f - \tau_0 = \mathcal{T}$ is the total duration of the trajectory \mathcal{S} , s_{-N} and s_N are the corresponding initial and final states and \mathcal{S}^* is the time reversal of \mathcal{S} . We recognize in this expression the *unnormalized* distributions of the λ and $(1-\lambda)$ -ensembles (see Eq. (10)). In the long-time limit $\mathcal{T} \rightarrow \infty$, the first factor appearing on the right side of Eq. (54) can be neglected (it is subdominant), leading to Eq. (52). For stochastic systems governed by time-dependent rate matrices, a transient result analogous to our Eq. (52) appears as Eq. (3.35) in Ref. [34]. However, to the best of our knowledge, the symmetry expressed by Eq. (52) has not previously been combined with the equivalence of ensembles (Eq. (14)) to arrive at our central result, Eq. (53), which has a simple and appealing physical interpretation.

Equation (53) is the stochastic version of Gallavotti's *conditional reversibility theorem* [23]. Gallavotti's theorem is derived in the context of deterministic dynamics and is stated in terms of "fluctuation patterns", a term that denotes the evolution of an arbitrary observable over a finite interval of time. However the basic content of the theorem is the same as that of Eq. (53): the probability to observe a system to behave in a particular manner when conditioning on one value of entropy production, is the same as the probability to observe the time-reversed behavior when conditioning on the opposite value. In particular, suppose we condition on the value $-\bar{\sigma}$, which is the opposite of the entropy production rate in the unbiased ensemble. Then Eq. (53) implies that the trajectories we will observe are statistically equivalent, under time-reversal, to the those generated by the original, unbiased dynamics, Eq. (1). In effect, in the ensemble conditioned on $\sigma = -\bar{\sigma}$, time will appear to be running backward. An analogous result has been obtained for the case of systems driven away from equilibrium by varying a parameter of the Hamiltonian [35, 36], rather than by equations of motion that violate detailed balance.

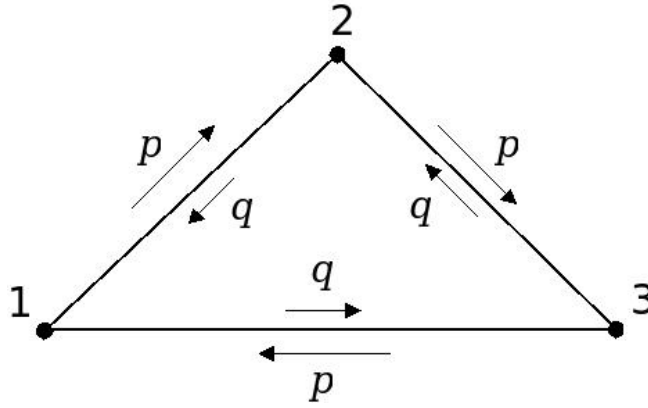


FIG. 1. Schematic representation of the transitions described by the matrix R . Clockwise transitions occur at an average rate p , and counterclockwise transitions at an average rate q .

IV. EXAMPLE

We now illustrate our results with an exactly solvable model. Consider the rate matrix

$$R = \begin{pmatrix} -1 & q & p \\ p & -1 & q \\ q & p & -1 \end{pmatrix}, \quad (55)$$

whose off-diagonal elements represent the transitions rates of a system with only three states, as shown schematically in Fig. 1. We impose $p + q = 1$ in order to fulfill the condition $\sum_i R_{ij} = 0$. (Note that q and p are *rates*, thus the condition $p + q = 1$ reflects our choice to set the escape rate from any state i to unity: $|R_{ii}| = 1$.)

We start by calculating the stationary currents, J_{ij} , and the infinite-time average entropy production rate, $\bar{\sigma}$, in the unbiased ensemble. The stationary distribution in this case is $|\pi\rangle = (1/3, 1/3, 1/3)^T$, from which we first obtain the currents,

$$J = \frac{1}{3} \begin{pmatrix} 0 & (q-p) & (p-q) \\ (p-q) & 0 & (q-p) \\ (q-p) & (p-q) & 0 \end{pmatrix}, \quad (56)$$

and then the average entropy production rate, using Eq. (9):

$$\bar{\sigma} = (p-q) \ln \left(\frac{p}{q} \right). \quad (57)$$

Analogously, the stationary currents $\tilde{J}_{ij}(\lambda)$ and the infinite-time average entropy production rate $\bar{\sigma}_\lambda$ for a biased ensemble are obtained from the transition rate matrix $\tilde{R}(\lambda)$. From Eq. (37), this matrix is constructed from the matrix $M(\lambda)$ (see Eq. (35)),

$$M(\lambda) = \begin{pmatrix} -1 & q \left(\frac{p}{q} \right)^\lambda & p \left(\frac{q}{p} \right)^\lambda \\ p \left(\frac{q}{p} \right)^\lambda & -1 & q \left(\frac{p}{q} \right)^\lambda \\ q \left(\frac{p}{q} \right)^\lambda & p \left(\frac{q}{p} \right)^\lambda & -1 \end{pmatrix} \equiv \begin{pmatrix} -1 & q_\lambda & p_\lambda \\ p_\lambda & -1 & q_\lambda \\ q_\lambda & p_\lambda & -1 \end{pmatrix}, \quad (58)$$

its largest eigenvalue,

$$\eta(\lambda) = -1 + q_\lambda + p_\lambda, \quad (59)$$

and the corresponding left-eigenvector,

$$\langle \omega(\lambda) | = (1, 1, 1). \quad (60)$$

Using Eqs. (55), (58), (59) and (60) in Eq. (37), we obtain

$$\tilde{R}(\lambda) = \begin{pmatrix} -(q_\lambda + p_\lambda) & q_\lambda & p_\lambda \\ p_\lambda & -(q_\lambda + p_\lambda) & q_\lambda \\ q_\lambda & p_\lambda & -(q_\lambda + p_\lambda) \end{pmatrix}. \quad (61)$$

The stationary distribution of (61) is $|\tilde{\pi}(\lambda)\rangle = (1/3, 1/3, 1/3)^T$, from which we obtain the stationary currents

$$\tilde{J}(\lambda) = \frac{1}{3} \begin{pmatrix} 0 & (q_\lambda - p_\lambda) & (p_\lambda - q_\lambda) \\ (p_\lambda - q_\lambda) & 0 & (q_\lambda - p_\lambda) \\ (q_\lambda - p_\lambda) & (p_\lambda - q_\lambda) & 0 \end{pmatrix}, \quad (62)$$

and the average entropy production rate,

$$\bar{\sigma}_\lambda = (p_\lambda - q_\lambda) \ln \left(\frac{p}{q} \right). \quad (63)$$

Finally, note that Eq. (58) implies the identity

$$q_\lambda = p_{1-\lambda}. \quad (64)$$

From Eq. 61 we see that in the biased ensemble, the system makes clockwise transitions at an average rate p_λ and counterclockwise transitions at a rate q_λ (compare with Fig. 1). Combining this observation with Eq. 64, we conclude that the dynamics in the λ -ensemble are the time-reversal of those in the $(1 - \lambda)$ -ensemble: the clockwise transition rates in one case become the counterclockwise rates in the other case. In particular, for $\lambda = 1$ the biased dynamics are obtained from the unbiased, $\lambda = 0$ dynamics by exchanging the roles of q and p in the rate matrix R . Moreover, Eqs. (47) and (48) are easily verified for an arbitrary value of λ , using Eqs. (62) - (64).

V. CONCLUSIONS

The equivalence of ensembles in the thermodynamic limit is a familiar and important concept in equilibrium statistical physics. In recent years analogous principles of equivalence have been developed for ensembles of trajectories representing systems away from thermal equilibrium [4–10], with the thermodynamic limit replaced by the long-time limit. In this paper we have applied these ideas to study discrete-state Markov processes that are conditioned on values of the time-averaged rate of entropy production. We first mapped the original Markov process onto a new, biased Markov process that describes the conditioned ensemble, as in Refs. [8–10]. We then used the properties of the biased transition rate matrix to establish our central result, which states that trajectories conditioned on opposite entropy production rates are related statistically by time-reversal. This extends Gallavotti’s conditional reversibility theorem [23], originally formulated for deterministic dynamics, to the case of stochastic dynamics.

We end by pointing out that the study of conditioned ensembles has often led to insights in nonequilibrium statistical physics. Most prominently, in two groundbreaking papers [37, 38] Onsager considered the spontaneous fluctuations of a system in equilibrium. By focusing on rare fluctuations that produce an “asymmetric distribution of energy” – or some other condition ordinarily associated with a system that is deliberately prepared away from equilibrium – Onsager was led to the *regression hypothesis* and the *reciprocal relations*, which lie at the foundations of linear response theory and the fluctuation-dissipation theorem. In recent years, Bertini *et al* [39–41] have developed *macroscopic fluctuation theory*, which builds on a large deviation-like formula for space-time fluctuations and considerably extends Onsager’s approach. Separately, Rahav and Jarzynski [42] have argued that when Onsager’s arguments are extended beyond the regime of linear response, they lead naturally to far-from-equilibrium *fluctuation theorems*. The common thread in these studies is a focus on spontaneous fluctuations conditioned on rare values of selected observables, such as energy distributions [37] or currents density profiles [41]. In the present paper, by considering rare fluctuations conditioned on entropy production, we have been led to our central result, which relates the sign of the conditioned entropy production to the direction of time’s arrow.

ACKNOWLEDGMENTS

M. B. acknowledges financial support from FAPESP (Brazil), Project No. 2012/07429-0, Unicamp/FAEPEX (Brazil), Grant No. 0031/15 and C. Jarzynski for his hospitality during the visit to University of Maryland, USA. C. J. acknowledges financial support from the National Science Foundation (USA) under Grant No. DMR-1206971.

Appendix: Results in the long-time limit when Q is defective

We discuss here how our results are modified when the matrix Q given by Eq. (18) is *defective* or non-diagonalizable. The main results of sections II and III rely basically on the asymptotic expressions

$$\langle \mathbf{1} | Q^n | \mathbf{p}(\tau_0) \rangle \simeq \mu^n \langle \mathbf{1} | \mathbf{u}_1 \rangle \langle \mathbf{v}_1 | \mathbf{p}(\tau_0) \rangle, \quad (\text{A.1a})$$

$$\langle \mathbf{e}_s | Q^n | \mathbf{p}(\tau_0) \rangle \simeq \mu^n \langle \mathbf{e}_s | \mathbf{u}_1 \rangle \langle \mathbf{v}_1 | \mathbf{p}(\tau_0) \rangle, \quad (\text{A.1b})$$

$$\langle \mathbf{1} | Q^n | \mathbf{e}_s \rangle \simeq \mu^n \langle \mathbf{1} | \mathbf{u}_1 \rangle \langle \mathbf{v}_1 | \mathbf{e}_s \rangle, \quad (\text{A.1c})$$

which are valid in the limit $n \gg 1$ since $|\alpha_k|/\mu < 1$ for $k \geq 2$. In what follows we show how to obtain such asymptotic expressions when Q is non-diagonalizable.

Under the assumptions spelled out in the first two paragraphs of Sec. II (but dropping the later assumption that Q is diagonalizable) the Perron-Frobenius theorem [29] guarantees that the eigenvalue $\alpha_1 = \mu$ with the largest real part is real and nondegenerate, hence we will continue to write $Q|\mathbf{u}_1\rangle = \mu|\mathbf{u}_1\rangle$ and $\langle\mathbf{v}_1|Q = \mu\langle\mathbf{v}_1|$. However, for those eigenvalues α_k whose algebraic and geometric multiplicities differ we have [28]

$$Q|\mathbf{u}_k^{(1)}\rangle = \alpha_k|\mathbf{u}_k^{(1)}\rangle, \quad (\text{A.2a})$$

$$Q|\mathbf{u}_k^{(\nu_k)}\rangle = \alpha_k|\mathbf{u}_k^{(\nu_k)}\rangle + |\mathbf{u}_k^{(\nu_k-1)}\rangle, \text{ for } 1 < \nu_k \leq \gamma(k), \quad (\text{A.2b})$$

and

$$\langle\mathbf{v}_k^{(\gamma(k))}|Q = \alpha_k\langle\mathbf{v}_k^{(\gamma(k))}|, \quad (\text{A.3a})$$

$$\langle\mathbf{v}_k^{(\nu_k)}|Q = \alpha_k\langle\mathbf{v}_k^{(\nu_k)}| + \langle\mathbf{v}_k^{(\nu_k+1)}|, \text{ for } 1 \leq \nu_k < \gamma(k), \quad (\text{A.3b})$$

where $\gamma(k)$ is the degeneracy of α_k . Equations (A.2) and (A.3) show that $|\mathbf{u}_k^{(1)}\rangle$ and $\langle\mathbf{v}_k^{(\gamma(k))}|$ are genuine eigenvectors. The $|\mathbf{u}_k^{(\nu_k)}\rangle$, for $1 < \nu_k \leq \gamma(k)$, and $\langle\mathbf{v}_k^{(\nu_k)}|$, for $1 \leq \nu_k < \gamma(k)$, are the so-called *generalized* eigenvectors. Together, genuine and generalized eigenvectors form a basis in which Q assumes the Jordan canonical form [28]. Therefore, the following relations apply

$$\langle\mathbf{v}_k^{(\nu_k)}|\mathbf{u}_j^{(\nu_j)}\rangle = \delta_{kj} \delta_{\nu_k \nu_j}, \quad (\text{A.4a})$$

$$\sum_k \sum_{\nu_k=1}^{\gamma(k)} |\mathbf{u}_k^{(\nu_k)}\rangle \langle\mathbf{v}_k^{(\nu_k)}| = \mathbb{1}, \quad (\text{A.4b})$$

which are the analogue of Eq. (20).

Equations (A.2) imply

$$Q^n |\mathbf{u}_k^{(\nu_k)}\rangle = \sum_{l=0}^{\nu_k-1} \binom{n}{l} \alpha_k^{n-l} |\mathbf{u}_k^{(\nu_k-l)}\rangle, \quad (\text{A.5})$$

for $n \geq \nu_k$ and $1 < \nu_k \leq \gamma(k)$. Analogously, Eqs. (A.3) imply

$$\langle\mathbf{v}_k^{(\nu_k)}|Q^n = \sum_{l=0}^{\gamma(k)-\nu_k} \binom{n}{l} \alpha_k^{n-l} \langle\mathbf{v}_k^{(\nu_k+l)}|, \quad (\text{A.6})$$

for $n > \gamma(k) - \nu_k$ and $1 \leq \nu_k < \gamma(k)$.

Expressions (A.5) and (A.6) allow us to find the dominant contribution of the matrix elements in the left-hand side of Eq. (A.1) when Q is defective. If we take for instance $\langle\mathbf{1}|Q^n|\mathbf{p}(\tau_0)\rangle$ and insert Eq. (A.4b), we obtain

$$\begin{aligned} \langle\mathbf{1}|Q^n|\mathbf{p}(\tau_0)\rangle &= \langle\mathbf{1}|Q^n \sum_k \sum_{\nu_k} |\mathbf{u}_k^{(\nu_k)}\rangle \langle\mathbf{v}_k^{(\nu_k)}| \mathbf{p}(\tau_0)\rangle \\ &= \langle\mathbf{1}|Q^n|\mathbf{u}_1\rangle \langle\mathbf{v}_1|\mathbf{p}(\tau_0)\rangle + \sum_{k=2} \sum_{\nu_k} \langle\mathbf{1}|Q^n|\mathbf{u}_k^{(\nu_k)}\rangle \langle\mathbf{v}_k^{(\nu_k)}|\mathbf{p}(\tau_0)\rangle \\ &= \mu^n \langle\mathbf{1}|\mathbf{u}_1\rangle \langle\mathbf{v}_1|\mathbf{p}(\tau_0)\rangle + \sum_{k=2} \sum_{\nu_k} \sum_{l=0}^{\nu_k-1} \binom{n}{l} \alpha_k^{n-l} \langle\mathbf{1}|\mathbf{u}_k^{(\nu_k-l)}\rangle \langle\mathbf{v}_k^{(\nu_k)}|\mathbf{p}(\tau_0)\rangle, \end{aligned} \quad (\text{A.7})$$

using Eq. (A.5) from the second to the third line. Analogously to Sec. II, we assume that $\langle \mathbf{1} | \mathbf{u}_1 \rangle$ and $\langle \mathbf{v}_1 | \mathbf{p}(\tau_0) \rangle$ are of the same magnitude of $\langle \mathbf{1} | \mathbf{u}_k^{(\nu_k-l)} \rangle$ and $\langle \mathbf{v}_k^{(\nu_k)} | \mathbf{p}(\tau_0) \rangle$, respectively. Thus, we recover Eq. (A.1a) from Eq. (A.7) if the following limit holds,

$$\lim_{n \rightarrow \infty} \binom{n}{l} \frac{\alpha_k^{n-l}}{\mu^n} = 0. \quad (\text{A.8})$$

In the limit $n \gg l$, we use an asymptotic expression for the binomial coefficient to obtain

$$\begin{aligned} \binom{n}{l} \left(\frac{|\alpha_k|}{\mu} \right)^n &= \binom{n}{l} e^{-n \ln(\mu/|\alpha_k|)} \\ &\simeq \frac{(n/l - 1/2)^l e^l}{\sqrt{2\pi l}} e^{-n \ln(\mu/|\alpha_k|)} \\ &< n^l e^{-n \ln(\mu/|\alpha_k|)}. \end{aligned} \quad (\text{A.9})$$

Since $\mu/|\alpha_k| > 1$ and l is always finite, we obtain

$$\lim_{n \rightarrow \infty} n^l e^{-n \ln(\mu/|\alpha_k|)} = 0, \quad (\text{A.10})$$

which, due to Eq. (A.9), implies Eq. (A.8). The same kind of analysis can be done for the other matrix elements of Eq. (A.1). In summary, Eqs. (A.1) still hold and our main results are valid when Q is defective.

-
- [1] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).
 - [2] H. Touchette, Europhys. Lett. **96**, 50010 (2011).
 - [3] H. Touchette, Phys. Rep. **478**, 1 (2009).
 - [4] R. M. L. Evans, Phys. Rev. Lett. **92**, 150601 (2004).
 - [5] R. M. L. Evans, J. Phys. A **38**, 293 (2005).
 - [6] C. Maes and K. Netočný, EPL **82**, 30003 (2008).
 - [7] R. M. L. Evans, Contemp. Phys. **51**, 413 (2010).
 - [8] R. L. Jack and P. Sollich, Prog. Theor. Phys. Suppl. **184**, 304 (2010).
 - [9] R. Chetrite and H. Touchette, Phys. Rev. Lett. **111**, 120601 (2013).
 - [10] R. Chetrite and H. Touchette, Ann. Henri Poincaré, **1** (2014).
 - [11] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
 - [12] S. Vaikuntanathan, T. R. Gingrich, and P. L. Geissler, Phys. Rev. E **89**, 062108 (2014).
 - [13] T. R. Gingrich, S. Vaikuntanathan, and P. L. Geissler, Phys. Rev. E **90**, 042123 (2014).
 - [14] M. Merolle, J. P. Garrahan, and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **102**, 10837 (2005).
 - [15] L. O. Hedges, R. L. Jack, J. P. Garrahan, and D. Chandler, Science **323**, 1309 (2009).
 - [16] D. Chandler and J. P. Garrahan, Ann. Rev. Chem. Phys. **61**, 191 (2010).
 - [17] C. Ates, B. Olmos, J. P. Garrahan, and I. Lesanovsky, Phys. Rev. A **85**, 043620 (2012).
 - [18] S. Genway, J. P. Garrahan, I. Lesanovsky, and A. D. Armour, Phys. Rev. E **85**, 051122 (2012).
 - [19] J. M. Hickey, S. Genway, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. A **86**, 063824 (2012).
 - [20] G. Verley, M. Esposito, T. Willaert, and C. Van den Broeck, Nature Communications **5**, 4721 (2014).
 - [21] T. R. Gingrich, G. Rotskoff, S. Vaikuntanathan, and P. L. Geissler, New Journal of Physics **16**, 102003 (2014).
 - [22] G. Verley, T. Willaert, C. Van den Broeck, and M. Esposito, Phys. Rev. E **90**, 052145 (2014).
 - [23] G. Gallavotti, Ann. de l'institut Henri Poincaré (A) **70**, 429 (1999).
 - [24] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
 - [25] P. Gaspard, J. Stat. Phys. **117**, 599 (2004).
 - [26] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
 - [27] A. Imparato and L. Peliti, J. Stat. Mech. **L02001** (2007).
 - [28] G. E. Shilov, *Linear Algebra* (Dover Publications, New York, 1977).
 - [29] E. Seneta, *Non-negative Matrices and Markov Chains* (Springer-Verlag, New York, 1981).
 - [30] J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart* (Springer, New York, 1984).
 - [31] J. Kurchan, J. Phys. A: Math. Gen. **31**, 37193729 (1998).
 - [32] J. G. Kemeny, J. L. Snell, and A. W. Knapp, *Denumerable Markov Chains* (Springer-Verlag, New York, 1976).
 - [33] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
 - [34] R. J. Harris and G. M. Schütz, J. Stat. Mech. **P07020** (2007).
 - [35] C. Jarzynski, Phys. Rev. E **73**, 046105 (2006).

- [36] C. Jarzynski, *Annu. Rev. Cond. Matt. Phys.* **2**, 329 (2011).
- [37] L. Onsager, *Phys. Rev.* **37**, 405 (1931).
- [38] L. Onsager, *Phys. Rev. B* **38**, 2265 (1931).
- [39] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Phys. Rev. Lett.* **87**, 040601 (2001).
- [40] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *J. Stat. Phys.* **107**, 635 (2002).
- [41] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Rev. Mod. Phys* **87**, 593 (2015).
- [42] S. Rahav and C. Jarzynski, *New Journal of Physics* **15**, 125029 (2013).