Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Classical statistical mechanics, equilibrium and non-equilibrium

Periodically driven jump processes conditioned on large deviations

Lydia Chabane[1], **Raphaël Chétrite**[2] **and Gatien Verley**[1]

[1] Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France
[2] Laboratoire J A Dieudonné, UMR CNRS 7351, Université de Nice Sophia Antipolis, Nice 06108, France
E-mail: lydia.chabane@universite-paris-saclay.fr

Received 7 November 2019
Accepted for publication 10 January 2020
Published 20 March 2020

Online at stacks.iop.org/JSTAT/2020/033208
https://doi.org/10.1088/1742-5468/ab74c4

Abstract. We study the fluctuations of systems modeled by time periodically driven Markov jump processes. We focus on observables defined through time-periodic functions of the system's states or transitions. Using large deviation theory, canonical biasing and Doob transform, we characterize the asymptotic fluctuations of such observables after a large number of periods by obtaining the Markov process that produces them. We show that this process, called driven process, is the optimizer under constraint of the large deviation function for occupation and jumps.

Keywords: large deviations in non-equilibrium systems, stochastic processes, fluctuation phenomena, stochastic thermodynamics

J. Stat. Mech. (2020) 033208

© 2020 IOP Publishing Ltd and SISSA Medialab srl

1742-5468/20/033208+25$33.00

Contents

1. Introduction **2**

2. Definitions and notations **4**
2.1. Periodically driven Markov jump processes 4
2.2. Observable, scaled cumulant generating function and tilted matrix 5
2.3. Spectral elements of the one-period propagator for the tilted process 6

3. Conditioning: building an effective process **7**
3.1. Microcanonical process 7
3.2. Canonical process 8
3.3. Driven process 9
3.4. Free-conditioning process 11

4. Variational representation of the driven process **12**

5. Illustration on a solvable modulated two-level system **16**

6. Conclusion **19**
Acknowledgment **20**
Appendix A. Definition of the Doob transform **20**
Appendix B. SCGF from the optimizer **21**
Appendix C. Definitions of the time-ordered exponential **22**
Appendix D. Properties of the time-ordered exponential **22**

References **24**

1. Introduction

Large deviation theory aims to predict the decay rate of a probability when increasing a parameter [1–3], e.g. number of realizations, time, system size. The stochastic variables we consider are physical observables characterizing the system of interest. Statistical properties of such observables follow then in an asymptotic limit. Large deviation theory is widely used in statistical physics [4, 5], both at equilibrium and out of equilibrium. At equilibrium, it gives a formal point of view on thermodynamics [6, 7] and thermodynamic potentials [8–13]. Within large deviation theory, state functions such as entropy or free energy are large deviation functions (LDFs) or cumulant generating functions (CGFs), and variational principles follow from a saddle point approximation. Out of equilibrium, the notion of state function disappears, but large deviation theory still brings a Legendre structure approaching that of equilibrium statistical mechanics [14]. This theory allows to compute transition probabilities between nonequilibrium states or statistics of observables such as activity, energy or matter currents. Examples include systems with noise-induced transitions [15–17], kinetically constrained models

J. Stat. Mech. (2020) 033208

[18–20], and chemical reaction networks for which a Lagrangian/Hamiltonian description has been derived in [21]. Algorithms computing numerically the probability of rare events have also been developed in the framework of large deviation theory [22–27]. More generally, the concept of large deviations appears in many fields such as dynamics of population [28, 29], finance [30] and bio-informatics [31].

Focusing on systems modeled by Markov jump processes, we are interested in the fluctuations of an observable $\boldsymbol{A}$ extensive in time that satisfies a large deviation principle. $\boldsymbol{A}$ may be a physical observable such as heat current, work or entropy production. We do not specify $\boldsymbol{A}$ as it exists no relevant set of extensive observables that can characterize any nonequilibrium steady state. However, we know from the large deviation perspective on statistical mechanics that the rare event statistics of a system in a given equilibrium state informs on all thermodynamic states of the same system. In the same line of idea, for nonequilibrium systems, a large deviation of $\boldsymbol{A}$ for a given process corresponds to a typical value of $\boldsymbol{A}$ for another process [14, 32, 33]. The question here is to determine the dynamics of the latter process, first to compare it with the former one, and second to be able to sample typical realizations leading to the chosen value $\boldsymbol{a}$ of observable $\boldsymbol{A}$. For this reason, we consider in the following a process submitted to a condition on a (rare) value $\boldsymbol{a}$ of observable $\boldsymbol{A}$. We then connect this conditioned process to an effective conditioning-free process taking $\boldsymbol{a}$ as a typical value. Such a connection has been done recently in the time-homogeneous case in [14, 34–36]. In this paper, we extend this result for periodically driven processes and for observables involving time-periodic functions.

This extension is motivated by the fact that many thermodynamic machines, including engines, operate via cycles or under periodic control. Such machines are experimentally studied nowadays at the fluctuating level [37–40]. Achieving a complete theory on the conditioning of those systems requires first to understand the fluctuations of periodically driven processes, which have attracted interest at the theoretical level [41–43]. Besides externally driven systems, spontaneous oscillations exist for systems in stationary nonequilibrium, see for instance the Brusselator model of chemical reactions [44, 45], or the three-state model studied in [46].

In this paper, we start in section 2 by defining our periodically driven Markov jump process and the conditioning observable $\boldsymbol{A}$. Then, we introduce the moment generating and the scaled cumulant generating functions (SCGF) for observable $\boldsymbol{A}$. Next, we express the SCGF in terms of the spectral elements associated with the one-period propagator for the generating function. In section 3, we address the problem of conditioning a Markov process on rare values of $\boldsymbol{A}$. This conditioned process corresponds to a microcanonical ensemble of trajectories, i.e. trajectories filtrated on the value of $\boldsymbol{A}$. However, this process has usually no Markov property. Like in equilibrium statistical mechanics, we use in section 3 the canonical ensemble of trajectories (by exponentially biasing each trajectory probability) to build a canonical process which is Markovian. In the limit of a large number of periods, we show that the canonical process becomes the so-called driven process that will appear later as an optimal process for which $\boldsymbol{A}$ converges in probability to the microcanonical value. Assuming a unique relation between the canonical bias and the conditioning value $\boldsymbol{a}$, the driven process defines the conditioning-free process which is asymptotically equivalent to the microcanonical process. In section 4, we show from a variational point of view that the driven process

J. Stat. Mech. (2020) 033208

is related to the optimizer of the 2.5 LDF of occupations and transition probabilities under the constraint $\boldsymbol{A} = \boldsymbol{a}$. Hence, the driven process is the most probable process that reproduces the fluctuation $\boldsymbol{a}$. This method is in clear analogy with Jaynes' maximum entropy principle in which entropy is replaced by the 2.5 LDF. In section 5, we conclude this paper by illustrating our results on a periodically modulated two level system conditioned on a current that is differently defined on each part of the period.

2. Definitions and notations

2.1. Periodically driven Markov jump processes

We consider a continuous-time Markov jump process defined on a finite state space [47, 48], $z(t)$ giving the state of the system at time t. The generator of this Markov process is denoted by $\boldsymbol{k}$ where $k_{xy} \equiv k_{xy}(t)$ for $x \neq y$ is the transition rate from state y to state x at time t (non-negative) and $k_{yy} \equiv -\lambda_y$ with

$$\lambda_y(t) \equiv \sum_{x \neq y} k_{xy}(t) \tag{1}$$

is the escape rate from state y at time t. We assume that the transition rates are time-periodic functions with period T, $\boldsymbol{k}(t+T) = \boldsymbol{k}(t)$, possibly with discontinuities. We denote by $\pi_x \equiv \pi_x(t)$ the probability to be in state x at time t. It satisfies the master equation

$$\frac{\partial \pi_x}{\partial t} = \sum_y k_{xy} \pi_y. \tag{2}$$

It remains normalized to 1 when it evolves according to the master equation since by construction $\sum_x k_{xy} = 0$, $\forall y$. We suppose that at time $t = nT$ and for $n \to \infty$, $\boldsymbol{\pi}$ reaches a periodic solution of the master equation $\boldsymbol{\pi}^{\mathrm{Tips}}$—where TiPS stands for *time periodic state*—i.e. $\boldsymbol{\pi}^{\mathrm{Tips}}(\tau + T) = \boldsymbol{\pi}^{\mathrm{Tips}}(\tau)$, $\forall \tau \in [0, T[$.

We call a path the succession of states visited by the system in addition to the knowledge of the times at which transitions occur. In the following, we denote a path by $[z]$ and assume that it starts at $t = 0$ and ends at $t = nT$, if not otherwise stated. We label $\{z_i\}_{i=0}^N$ the visited states and $\{t_i\}_{i=0}^N$ the times at which the system jumps such that

$$z(t) = z_i \qquad \text{for} \qquad t_i \leqslant t < t_{i+1}. \tag{3}$$

Time $t_0 = 0$ is the initial time and t_N is the last jump time before the final time nT. The path probability $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$ of path $[z]$ is given by

$$\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z] = \pi_{z_0}(0) \exp\left[\sum_{i=0}^{N-1} \ln\left(k_{z_{i+1},z_i}(t_{i+1})\right) - \int_0^{nT} \lambda_{z(t)}(t) \mathrm{d}t\right], \tag{4}$$

where the index $\boldsymbol{\pi}(0)$ refers to the initial state probability.

J. Stat. Mech. (2020) 033208

2.2. Observable, scaled cumulant generating function and tilted matrix

We are now interested in the fluctuations of an observable $\boldsymbol{A}_{nT} \equiv \boldsymbol{A}_{nT}[z]$ that is a real functional of the paths up to the final time nT. For sake of generality, we consider the following two-components observable $\boldsymbol{A}_t[z]$ on the shorter time interval $[0, t]$ defined by

$$\boldsymbol{A}_t[z] \equiv \frac{1}{t} \begin{pmatrix} \displaystyle\sum_{i=0}^{N-1} g_{z_{i+1},z_i}(t_{i+1})\theta(t - t_{i+1}) \\ \displaystyle\int_0^t \mathrm{d}\tau f_{z(\tau)}(\tau) \end{pmatrix} \tag{5}$$

where $\theta(t)$ is the Heavyside function, $\boldsymbol{g}$ and $\boldsymbol{f}$ are time periodic functions with period T, i.e. $\boldsymbol{g}(\tau + T) = \boldsymbol{g}(\tau)$ and $\boldsymbol{f}(\tau + T) = \boldsymbol{f}(\tau)$, $\forall \tau \in \mathbb{R}$. When specifying the components of matrices $\boldsymbol{g}$ and vector $\boldsymbol{f}$, observable $\boldsymbol{A}_{nT}[z]$ may represent physical quantities. For instance, it will be the number of jumps per unit time if $g_{xy} = 1$ and $f_y = 0$ for all x and y, and the occupation time in state x if $g_{y,z} = 0$ and $f_y = \delta_{xy}$ with δ the Kronecker delta.

We assume that $\boldsymbol{A}_{nT}$ satisfies the large deviation principle

$$P(\boldsymbol{a}) \underset{n\to\infty}{\sim} \mathrm{e}^{-nTI(\boldsymbol{a})}, \tag{6}$$

where $P(\boldsymbol{a})$ is the probability of the rare event $\{\boldsymbol{A}_{nT}[z] = \mathbf{a} \mid \mathbf{a} \in \mathbb{R}^2\}$ and $I(\boldsymbol{a})$ its associated LDF or rate function. The LDF describes the exponential decay with time of the probability that $\boldsymbol{A}_{nT}$ takes a value at an arbitrary distance from its typical one. In order to study the fluctuations of $\boldsymbol{A}_{nT}$ in the long-time limit, we introduce the generating function for $\boldsymbol{A}_t$

$$G(\boldsymbol{\gamma}, t) \equiv \mathbb{E}_{\boldsymbol{\pi}(0)} \left[\mathrm{e}^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\right], \tag{7}$$

where $\mathbb{E}_{\boldsymbol{\pi}(0)}[\cdot]$ is the path average on $[z]$ with initial probability $\boldsymbol{\pi}(0)$. The central dot $\cdot$ stands for the scalar product and vector $\boldsymbol{\gamma} = \begin{pmatrix}\gamma_1 & \gamma_2\end{pmatrix}^{\dagger}$ is the conjugate variable of $\boldsymbol{A}_t$. The generating function at time t imposing the final state $z(t) = x$ writes

$$G_x(\boldsymbol{\gamma}, t) \equiv \mathbb{E}_{\boldsymbol{\pi}(0)} \left[\mathrm{e}^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\delta_{x,z(t)}\right]. \tag{8}$$

The SCGF for $\boldsymbol{A}_{nT}$ is defined by

$$\phi(\boldsymbol{\gamma}) \equiv \lim_{n\to\infty} \frac{1}{nT} \ln G(\boldsymbol{\gamma}, nT). \tag{9}$$

The conditioned generating function G_x satisfies the ordinary differential equation [43, 49–51]

$$\partial_t G_x(\boldsymbol{\gamma}, t) = \sum_y \kappa_{xy}(\boldsymbol{\gamma}, t) G_y(\boldsymbol{\gamma}, t), \tag{10}$$

where we have introduced the tilted (or dressed) operator $\boldsymbol{\kappa}$ of components

$$\kappa_{xy}(\boldsymbol{\gamma}, t) \equiv \begin{cases} k_{xy}(t)\mathrm{e}^{\gamma_1 g_{xy}(t)} & \text{if } x \neq y, \\ -\lambda_x(t) + \gamma_2 f_x(t) & \text{if } x = y. \end{cases} \tag{11}$$

Notice that by definition $\boldsymbol{\kappa}(\boldsymbol{\gamma}, \tau + T) = \boldsymbol{\kappa}(\boldsymbol{\gamma}, \tau)$, $\forall \tau \in \mathbb{R}$. In the following, we keep in mind that $\boldsymbol{\kappa}$ depends on $\boldsymbol{\gamma}$ and drop $\boldsymbol{\gamma}$ in the notations for clarity. The tilted matrix

J. Stat. Mech. (2020) 033208

can be seen as the generator of a new process, called the tilted process, that contains information on the large deviations of $\boldsymbol{A}_{nT}$, but that is not norm conserving since $\sum_x \kappa_{xy} \neq 0$, $\forall y$. We formally solve equation (10) with initial condition $\boldsymbol{G}(\boldsymbol{\gamma}, 0) = \boldsymbol{\pi}(0)$ writing

$$\boldsymbol{G}(\boldsymbol{\gamma}, t) = \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t, 0)\boldsymbol{\pi}(0), \tag{12}$$

in terms of the propagator

$$\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t, t_0) \equiv \overleftarrow{\exp} \int_{t_0}^{t} \boldsymbol{\kappa}(t')\,\mathrm{d}t', \tag{13}$$

involving the time-ordered exponential $\overleftarrow{\exp}$, see appendix C. $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t, t_0)$ is the unique solution of the initial matrix value problem $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{\kappa}(t)\boldsymbol{X}(t)$, with $\boldsymbol{X}(0) = \mathbb{1}$ the identity matrix in the state space.

2.3. Spectral elements of the one-period propagator for the tilted process

From now on, we assume that the final time is always nT and omit the subscript nT for our generic observable $\boldsymbol{A}[z] \equiv \boldsymbol{A}_{nT}[z]$. This observable evaluated for a stochastic process $z(t)$ on $[0, nT]$ becomes the random variable $\boldsymbol{A}$. In this section, we relate generating functions for $\boldsymbol{A}$ to the spectrum and eigenspace of the propagator $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T, 0)$, giving them a physical interpretation.

Let ρ_T be the largest (dominant) eigenvalue of $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T, 0)$ (unicity is ensured by Perron–Frobenius theorem), $\boldsymbol{r}_T$ its right (column) eigenvector and $\boldsymbol{l}_T$ its left (row) eigenvectors:

$$\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T, 0)\boldsymbol{r}_T = \rho_T\,\boldsymbol{r}_T, \tag{14}$$

$$\boldsymbol{l}_T\,\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T, 0) = \rho_T\,\boldsymbol{l}_T. \tag{15}$$

The eigenvectors $\boldsymbol{l}_T$ and $\boldsymbol{r}_T$ can be chosen up to a multiplicative constant that we set by imposing

$$\mathbf{1} \cdot \boldsymbol{r}_T = 1, \tag{16}$$

$$\boldsymbol{l}_T \cdot \boldsymbol{r}_T = 1, \tag{17}$$

where $\mathbf{1}$ is the vector whose components are all 1. Furthermore, we remark that $\boldsymbol{l}_T \cdot \boldsymbol{\pi}(0) < \infty$ since the state space is finite.

We now make a connection between the spectral elements of the propagator and the large deviations of $\boldsymbol{A}$. From equations (8) and (12), property D.5 of appendix D and using the periodicity of $\boldsymbol{\kappa}$, the generating functions G and G_x at time nT write

$$G(nT) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\right] = \mathbf{1} \cdot \left(\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T, 0)^n \boldsymbol{\pi}(0)\right) = \sum_{x,y} \left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T, 0)^n\right]_{xy} \pi_y(0), \tag{18}$$

J. Stat. Mech. (2020) 033208

$$G_x(nT) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\delta_{x,z(nT)}\right] = \sum_y \left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)^n\right]_{xy}\pi_y(0). \tag{19}$$

The asymptotic expansion of $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(nT,0)$ at large n is given by

$$\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0)^n \underset{n\to\infty}{\approx} (\rho_T)^n\,\boldsymbol{r}_T\,\boldsymbol{l}_T. \tag{20}$$

With equations (16)–(18), it yields

$$\phi(\boldsymbol{\gamma}) = \lim_{n\to\infty}\frac{1}{nT}\ln\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\right] = \frac{1}{T}\ln\rho_T. \tag{21}$$

The SCGF ϕ is proportional to the logarithm of the largest eigenvalue of the single-period propagator $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0)$ [43, 51]. Similarly, combining equations (18)–(20) and using equations (16) and (17), we find

$$\lim_{n\to\infty}\mathrm{e}^{-nT\phi}\mathbb{E}_{x_0}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\right] = (l_T)_{x_0}, \tag{22}$$

$$\lim_{n\to\infty}\frac{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\delta_{z(nT),x}\right]}{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\right]} = (r_T)_x, \tag{23}$$

where $\mathbb{E}_{x_0}$ is the path average over $[z]$ with a Kronecker delta centered on x_0 as initial probability. Hence, equations (22) and (23) allow to write the eigenvectors of the propagator in terms of path averages. Hence, we have extended for periodically driven processes the results of [35] that hold in the time-homogeneous case.

3. Conditioning: building an effective process

We saw in the previous section that the propagator based on the tilted matrix allows to describe the large deviations of $\boldsymbol{A}$. We are now interested in conditioning our original Markov process by filtering the ensemble of paths to select those leading to a chosen value of $\boldsymbol{A}$. This defines the so-called microcanonical process for which we aim to find an equivalent Markov process in the long-time limit. Some results of this section are stated in section 8.8 of [52].

3.1. Microcanonical process

The process $z(t)$ conditioned on the event $\{\boldsymbol{A}[z] = \mathbf{a} \mid \mathbf{a}\in\mathbb{R}^2\}$ is described by the microcanonical path probability [35]

$$\mathbb{P}^{\mathrm{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}\left[z \mid \boldsymbol{A}[z] = \mathbf{a}\right]. \tag{24}$$

In general, there is no Markov generator that can exactly generate this microcanonical ensemble of paths. Yet, there is another process called the canonical process that is Markovian and that has the interesting property to be asymptotically equivalent (in a way to be defined later) to the microcanonical process [35, 53].

J. Stat. Mech. (2020) 033208

3.2. Canonical process

The canonical path probability is connected to the original process by an exponential tilting of the path probability $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$ [35]:

$$\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z] \equiv \frac{\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]}{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\right]} = \frac{\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]}{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\right]}, \tag{25}$$

where

$$\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]} \tag{26}$$

is obtained by inserting equation (11) in the definition of a path probability of equation (4). This path probability is a natural generalization at the path level of the equilibrium probability in the canonical ensemble. This definition has already been used in many articles, for instance for the simulation of transition paths associated with glassy systems [54, 55].

The canonical process is Markovian: it is generated by a norm conserving Markov generator that we derive in the following. To do so, we look for a generator $\boldsymbol{\mathcal{K}}^n$ that satisfies

$$\mathbb{P}_{\boldsymbol{\mathcal{K}}^n,\boldsymbol{\pi}'(0)}[z] \equiv \mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z], \tag{27}$$

where $\boldsymbol{\pi}'(0)$ is an initial probability that may be different from $\boldsymbol{\pi}(0)$. Moreover, we want $\boldsymbol{\mathcal{K}}^n$ to generate a process that conserves probability, hence we look for a vector $\boldsymbol{C}^n$ such that

$$\boldsymbol{\mathcal{K}}^n \equiv \boldsymbol{\kappa}^{\boldsymbol{C}^n} \equiv \mathcal{D}(\boldsymbol{C}^n)\boldsymbol{\kappa}\,\mathcal{D}(\boldsymbol{C}^n)^{-1} - \mathcal{D}(\boldsymbol{C}^n)^{-1}\mathcal{D}(\boldsymbol{C}^n\boldsymbol{\kappa}) \tag{28}$$

is built from a Doob transform as defined in appendix A. Looking at the path probability in equation (A.3) obtained from a Doob transform, $\boldsymbol{C}^n \equiv \boldsymbol{C}^n(t)$ should be chosen as the solution of

$$\begin{cases} \dot{\boldsymbol{C}}^n = -\boldsymbol{C}^n\boldsymbol{\kappa}, \\ \boldsymbol{C}^n(nT) = \mathbf{1}, \end{cases} \tag{29}$$

such that the time extensive term in the exponential of equation (A.3) vanishes. Equivalently, using property D.3 of appendix D, $\boldsymbol{C}^n$ writes

$$\boldsymbol{C}^n(t) = \mathbf{1}\,\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(nT,t). \tag{30}$$

Inserting equation (28) in equation (A.3), we obtain that the path probability associated with $\boldsymbol{\kappa}^{\boldsymbol{C}^n}$ is given by

$$\mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]\,(C^n_{x_0})^{-1}(0), \tag{31}$$

or equivalently

$$\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n},\,\boldsymbol{\pi}(0)\circ\boldsymbol{C}^n(0)}[z], \tag{32}$$

where $\circ$ is the Hadamard product: $(\boldsymbol{u}\circ\boldsymbol{v})_x \equiv u_x v_x$. From equations (18) and (30), we remark that the generating function can be expressed in terms of $\boldsymbol{C}^n$ as

J. Stat. Mech. (2020) 033208

$$G(nT) = \boldsymbol{C}^n(0) \cdot \boldsymbol{\pi}(0), \tag{33}$$

implying

$$\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n}, \frac{\boldsymbol{C}^n(0)\circ\boldsymbol{\pi}(0)}{\boldsymbol{C}^n(0)\cdot\boldsymbol{\pi}(0)}}[z] \tag{34}$$

from equations (25) and (32). In other words, the canonical path probability is associated with the time-dependent generator $\boldsymbol{\mathcal{K}}^n = \boldsymbol{\kappa}^{\boldsymbol{C}^n}$ for the initial probability $\boldsymbol{C}^n(0) \circ \boldsymbol{\pi}(0)/[\boldsymbol{C}^n(0) \cdot \boldsymbol{\pi}(0)]$. This shows that the canonical process has a corresponding Markov generator and can thus be considered as a Markov process. Note that this generator depends explicitly on the number of periods n.

In the next section, we focus on the asymptotic dynamics in the limit $n \to \infty$ by considering the process towards which the canonical process converges at long time.

3.3. Driven process

The driven process is defined as the limit of the canonical process as $n \to \infty$ [35]. Since the canonical process comes from the Doob transform of the tilted operator using $\boldsymbol{C}^n$, the driven process will be built similarly. In the limit $n \to \infty$, using equations (16) and (20), we find that $\boldsymbol{C}^n(\tau)$ for $\tau \in [0, T[$ is given asymptotically by

$$\boldsymbol{C}^n(\tau) = \mathbf{1}\ \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0)^n \left[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(\tau,0)\right]^{-1} \underset{n\to\infty}{\sim} (\rho_T)^n \boldsymbol{l}_T \left[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(\tau,0)\right]^{-1}. \tag{35}$$

Note that Abel–Jacobi–Liouville formula ensures the invertibility of $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(\tau,0)$. Since scalar constants play no role in the Doob transform, we introduce the function of time $\boldsymbol{l} \equiv \boldsymbol{l}(\tau)$:

$$\boldsymbol{l}(\tau) \equiv \boldsymbol{l}_T \left[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(\tau,0)\right]^{-1}, \tag{36}$$

that is by construction the solution of

$$\begin{cases} \dot{\boldsymbol{l}} = -\boldsymbol{l}\boldsymbol{\kappa}, \\ \boldsymbol{l}(0) = \boldsymbol{l}_T. \end{cases} \tag{37}$$

Using equation (15) and the periodicity of $\boldsymbol{\kappa}$, the vector $\boldsymbol{l}$ satifies

$$\boldsymbol{l}(\tau + T) = \rho_T^{-1} \boldsymbol{l}(\tau). \tag{38}$$

We define the Markov generator $\boldsymbol{K} \equiv \boldsymbol{K}(\boldsymbol{\gamma}, \tau)$ of the driven process at all time τ by the Doob transform of the tilted matrix $\boldsymbol{\kappa}$ associated with vector $\boldsymbol{l}$:

$$\boldsymbol{K} \equiv \boldsymbol{\kappa}^{\boldsymbol{l}} = \mathcal{D}(\boldsymbol{l})\boldsymbol{\kappa}\,\mathcal{D}(\boldsymbol{l})^{-1} - \mathcal{D}(\boldsymbol{l})^{-1}\mathcal{D}(\boldsymbol{l}\boldsymbol{\kappa}). \tag{39}$$

Note that the positivity of $\boldsymbol{l}(t)$ at all t is ensured by the positivity of $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t,0)$ and Perron–Frobenius theorem (which ensures the positivity of $\boldsymbol{l}_T$). For time-homogeneous processes, i.e. for time-independent generators $\boldsymbol{k}(t) = \boldsymbol{k}$ and $\boldsymbol{\kappa}(t) = \boldsymbol{\kappa}$, $\boldsymbol{l}_T$ becomes T-independent and is the left-eigenvector of $\mathrm{e}^{T\boldsymbol{\kappa}}$ (and of $\boldsymbol{\kappa}$) associated with its dominant eigenvalue ρ_T (respectively $\frac{1}{T}\ln\rho_T$). The function $\boldsymbol{l}(\tau)$ in (36) becomes $\boldsymbol{l}(\tau) = \boldsymbol{l}_T \mathrm{e}^{-\tau\boldsymbol{\kappa}} = \mathrm{e}^{-\frac{\tau}{T}\ln\rho_T}\boldsymbol{l}_T$. Since the exponential term disappear in the Doob transform, the driven process is the

J. Stat. Mech. (2020) 033208

Doob transform of the tilted matrix with respect to $\boldsymbol{l}_T$, which is consistent with the results of [35].

From equations (28) and (35), we see that the generator of the driven process is given by the limit of the canonical transition matrix as $n \to \infty$:

$$\lim_{n\to\infty} \boldsymbol{\mathcal{K}}^n(\boldsymbol{\gamma},\tau) = \boldsymbol{\kappa}^{\boldsymbol{l}}(\boldsymbol{\gamma},\tau) = \boldsymbol{K}(\boldsymbol{\gamma},\tau). \tag{40}$$

One interesting property of $\boldsymbol{K}$ is its periodicity. Indeed, from equation (38) and the periodicity of $\boldsymbol{\kappa}$, we have $\boldsymbol{K}(\boldsymbol{\gamma},\tau+T) = \boldsymbol{K}(\boldsymbol{\gamma},\tau),\ \forall\tau\in\mathbb{R}$.

In the following, we show that the driven and canonical path probabilities are asymptotically equivalent. Two paths $\mathbb{P}_n$ and $\mathbb{Q}_n$ are said to be logarithmically equivalent if $\lim_{n\to\infty}\frac{1}{n}\ln\frac{\mathbb{P}_n}{\mathbb{Q}_n}$, and we denote it $\mathbb{P}_n \asymp \mathbb{Q}_n$. In this case, if an observable satisfies a large deviation principle with respect to $\mathbb{P}_n$ and $\mathbb{Q}_n$, then the corresponding LDFs vanish at the same values. This means that, in case of logarithmic equivalence, this observable takes the same typical values with respect to both paths in the limit $n\to\infty$ [35, 53]. Using equations (21), (37), (39) and (A.3), the path probability of the driven process writes

$$\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]\, l_{z_{nT}}(0)\, \mathrm{e}^{-nT\phi}\, l_{z_0}^{-1}(0). \tag{41}$$

Using the definitions of the canonical path probability (25) and driven path probability (41), we get

$$\frac{\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z]}{\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]} = l_{z_{nT}}(0)\, \mathrm{e}^{-nT\phi}\, l_{z_0}(0)\, \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}[z]}\right]. \tag{42}$$

Hence, using the definition of the SCGF (9), we finally find:

$$\lim_{n\to\infty}\frac{1}{nT}\ln\frac{\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z]}{\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]} = 0. \tag{43}$$

The driven path probability and the canonical path probability are then logarithmically equivalent:

$$\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z] \asymp \mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]. \tag{44}$$

Finally, we remark that the TiPS probability for the driven process can be obtained from the solution of the initial value problem of equation (37) for $\boldsymbol{l}$ and the intial value problem for $\boldsymbol{r}\equiv\boldsymbol{r}(\tau)$

$$\begin{cases}\dot{\boldsymbol{r}} = \boldsymbol{\kappa}\boldsymbol{r},\\ \boldsymbol{r}(0) = \boldsymbol{r}_T,\end{cases} \tag{45}$$

or alternatively

$$\boldsymbol{r}(t) \equiv \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t,0)\boldsymbol{r}_T. \tag{46}$$

Using equation (14) and the periodicity of $\boldsymbol{\kappa}$, the vector $\boldsymbol{r}$ satisfies:

$$\boldsymbol{r}(t+T) = \rho_T\boldsymbol{r}(t). \tag{47}$$

J. Stat. Mech. (2020) 033208

The TiPS probability of the driven process $\boldsymbol{\mu} \equiv \boldsymbol{\mu}(t)$, defined as the T-periodic solution of the master equation:

$$\begin{cases} \frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \boldsymbol{K}\boldsymbol{\mu} \\ \boldsymbol{\mu}(0) = \boldsymbol{\mu}(T), \end{cases} \tag{48}$$

writes in terms of the vectors $\boldsymbol{l}$ and $\boldsymbol{r}$

$$\boldsymbol{\mu}(t) = \boldsymbol{l}(t) \circ \boldsymbol{r}(t). \tag{49}$$

Indeed, equations (37), (39) and (45) yield for all x

$$\sum_y K_{xy}(l_y r_y) = \sum_y \left\{ l_x \kappa_{xy} l_y^{-1} l_y r_y - l_x^{-1} l_y \kappa_{yx} l_x r_x \right\} \tag{50}$$

$$= l_x \dot{r}_x + \dot{l}_x r_x \tag{51}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}(l_x r_x), \tag{52}$$

while equations (38) and (47) lead to

$$\boldsymbol{l}(0) \circ \boldsymbol{r}(0) = \boldsymbol{l}(T) \circ \boldsymbol{r}(T), \tag{53}$$

which proves that $\boldsymbol{\mu}$ is the solution of equation (48). Notice that our normalization choice in equation (17) ensures the normalization of $\boldsymbol{\mu}(0)$. We emphasize that in this section we have essentially extended for the periodic case the results of [35] for the time-homogeneous case.

3.4. Free-conditioning process

We saw that the canonical process tends to the driven process in the long-time limit. We still need to obtain the Markov process that is equivalent at large time to the microcanonical process for which $\{\boldsymbol{A}[z] = \mathbf{a} \mid \mathbf{a} \in \mathbb{R}^2\}$. This process is called the *free-conditioning process*.

From [53], the canonical and microcanonical path probabilities (thought respectively as biased and conditioned path ensembles based on $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$ and observable $\boldsymbol{A}$) are logarithmically equivalent if the LDF I is convex at $\boldsymbol{a}$. In this case, and assuming that I is differentiable for simplicity, the equivalence holds for $\boldsymbol{\gamma} = \nabla I(\boldsymbol{a})$, where $\nabla I(\boldsymbol{a})$ is the gradient of I evaluated at $\boldsymbol{a}$. Mathematically, this writes:

$$\mathbb{P}^{\mathrm{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] \asymp \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z] \,\big|_{\boldsymbol{\gamma}=\nabla I(\boldsymbol{a})} \,. \tag{54}$$

When combined with the logarithmic equivalence between the driven and canonical path probabilities of equation (44), we find that the free-conditioning process is the driven process for $\boldsymbol{\gamma} = \nabla I(\boldsymbol{a})$. Mathematically, this writes:

$$\mathbb{P}^{\mathrm{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] \asymp \mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z] \,\big|_{\boldsymbol{\gamma}=\nabla I(\boldsymbol{a})} \,. \tag{55}$$

Notice that if I is not convex at $\boldsymbol{a}$, there is no Markov process equivalent to the microcanonical process.

J. Stat. Mech. (2020) 033208

4. Variational representation of the driven process

In this section, we derive the driven process from a variational approach. This route requires to determine the functional to be minimized and playing the role of entropy in Jaynes' maximum entropy principle of statistical mechanics. This functional is the level 2.5 LDF for occupations and transition probabilities. We find that the driven process is the 'most probable' process for which observable $\boldsymbol{A}$ takes asymptotically a chosen value.

The *occupation density* $p_x^n(\tau)[z]$ at phase $\tau \in [0, T[$ is the path functional

$$p_x^n(\tau)[z] = \frac{1}{n}\sum_{m=0}^{n-1} \delta_{x,z(\tau+mT)} \tag{56}$$

that counts the fraction of time the system has been in state x at phase τ of each period along the path $[z]$. The occupation density is a positive vector of normalization 1 that converges to $\boldsymbol{\pi}^{\text{Tips}}$, when considering the process of generator $\boldsymbol{k}$. The *empirical transition probability* $\omega_{xy}^n(\tau)[z]$

$$\omega_{xy}^n(\tau)[z] = \frac{1}{n}\sum_{m=0}^{n-1}\frac{1}{\mathrm{d}\tau}\sum_{s\in[\tau,\tau+\mathrm{d}\tau]} \delta_{y,z(s^-+mT)}\delta_{x,z(s^++mT)}, \tag{57}$$

with $\mathrm{d}\tau$ an infinitesimal time, measures the number of transitions $y \to x$ per unit of time at phase τ, or more precisely during $[\tau, \tau + d\tau]$. When considering the process of generator $\boldsymbol{k}$, $\boldsymbol{\omega}$ converges to $\boldsymbol{k}\circ\boldsymbol{\pi}^{\text{Tips}}$, with $\left[\boldsymbol{k}\circ\boldsymbol{\pi}^{\text{Tips}}\right]_{xy} \equiv k_{xy}\pi_y^{\text{Tips}}$. We can rewrite the conditioning observable $\boldsymbol{A}$ of equation (5) in terms of $\boldsymbol{p}^n[z]$ and $\boldsymbol{\omega}^n[z]$ using the periodicity of $\boldsymbol{f}$ and $\boldsymbol{g}$:

$$\boldsymbol{A}\left(\boldsymbol{\omega}^n[z], \boldsymbol{p}^n[z]\right) = \begin{pmatrix} A_1(\boldsymbol{\omega}^n[z]) \\ A_2(\boldsymbol{p}^n[z]) \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{T}\int_0^T \mathrm{d}\tau \sum_{x,y\neq x} \omega_{xy}^n[z](\tau)\ g_{xy}(\tau) \\ \frac{1}{T}\int_0^T \mathrm{d}\tau \sum_x p_x^n[z](\tau)\ f_x(\tau) \end{pmatrix}. \tag{58}$$

At long time nT, the probability to observe the occupation $\boldsymbol{p}^n[z] = \boldsymbol{p}$ and the empirical transition probability $\boldsymbol{\omega}^n[z] = \boldsymbol{\omega}$ satifies a large deviation principle:

$$P(\boldsymbol{\omega}, \boldsymbol{p}) \underset{n\to\infty}{\sim} \mathrm{e}^{-nTI_{2.5}(\boldsymbol{\omega},\boldsymbol{p})}, \tag{59}$$

where the 2.5 LDF is given by [56]

$$I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p}) = \sum_{y,x\neq y}\frac{1}{T}\int_0^T \mathrm{d}\tau \left[p_y(\tau)\ \left(k_{xy}(\tau) - \frac{\omega_{xy}(\tau)}{p_y(\tau)}\right) + \omega_{xy}(\tau)\ln\frac{\omega_{xy}(\tau)}{k_{xy}(\tau)\,p_y(\tau)}\right], \tag{60}$$

see [57] for the equivalent LDF in the diffusive case and [58–61] in the time-homogeneous case. Equation (60) holds only for $\boldsymbol{p}(0) = \boldsymbol{p}(T)$ and $\boldsymbol{\omega}(0) = \boldsymbol{\omega}(T)$, and for conservative transition probabilities with $\dot{p}_x(\tau) = \sum_y(\omega_{xy}(\tau) - \omega_{yx}(\tau))$, $\forall x$, and normalized occupations with $\sum_y p_y(\tau) = 1$, $\forall y$, otherwise $I_{2.5}$ is infinite. Notice that $I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p})$ vanishes for $\boldsymbol{p} = \boldsymbol{\pi}^{\text{Tips}}$ and $\boldsymbol{\omega} = \boldsymbol{k}\circ\boldsymbol{\pi}^{\text{Tips}}$. Hence, without conditioning, $\boldsymbol{A}$ converges to $\boldsymbol{A}(\boldsymbol{k}\circ\boldsymbol{\pi}^{\text{Tips}}, \boldsymbol{\pi}^{\text{Tips}})$ as $n \to \infty$.

J. Stat. Mech. (2020) 033208

As before, we are interested in conditioning our process on the event $\{\boldsymbol{A}(\boldsymbol{\omega},\boldsymbol{p})=\boldsymbol{a} \mid \boldsymbol{a}\in\mathbb{R}^2\}$. We look for the most probable pair $(\boldsymbol{\omega},\boldsymbol{p})$ compatible with $\boldsymbol{A}(\boldsymbol{\omega},\boldsymbol{p})=\boldsymbol{a}$. This pair coincides with the typical value of $(\boldsymbol{\omega}^n[z],\boldsymbol{p}^n[z])$ under the microcanonical path probability [5, 36]. It is obtained by minimizing the 2.5 LDF under the following constraints

- **C0:** $\boldsymbol{A}(\boldsymbol{\omega},\boldsymbol{p})=\boldsymbol{a}$,
- **C1:** $\sum_y p_y(\tau)=1, \forall\tau\in[0,T[$,
- **C2:** $\dot{p}_x(\tau)=\sum_y(\omega_{xy}(\tau)-\omega_{yx}(\tau)), \forall x$ and $\forall\tau\in[0,T[$,
- **C3:** $\boldsymbol{p}(T)=\boldsymbol{p}(0)$,
- **C4:** $\boldsymbol{\omega}(T)=\boldsymbol{\omega}(0)$.

This optimization problem amounts to computing the LDF of $\boldsymbol{A}$ and writes mathematically

$$I(\boldsymbol{a})=\inf_{\boldsymbol{p},\boldsymbol{\omega}|\{\mathbf{Ci}\}_{i=0}^4}\Big\{I_{2.5}(\boldsymbol{\omega},\boldsymbol{p})\Big\}. \tag{61}$$

Equation (61) is known as the *contraction principle*. Intuitively, the optimizer $(\boldsymbol{\omega_a},\boldsymbol{p_a})$ is expected to be associated with the generator of the conditioning-free process or similarly the microcanonical process as $n\to\infty$. Equivalently, we can instead compute the SCGF given by the Legendre tranform of the LDF in equation (61):

$$\phi(\boldsymbol{\gamma})=\sup_{\boldsymbol{p},\boldsymbol{\omega}|\{\mathbf{Ci}\}_{i=1}^4}\left\{\boldsymbol{\gamma}\cdot\boldsymbol{A}(\boldsymbol{\omega},\boldsymbol{p})-I_{2.5}(\boldsymbol{\omega},\boldsymbol{p})\right\}. \tag{62}$$

The optimizer $(\boldsymbol{\omega_\gamma},\boldsymbol{p_\gamma})$ of equation (62) is the typical value of $(\boldsymbol{\omega}^n[z],\boldsymbol{p}^n[z])$ under the canonical path probability [5, 36]. It is expected to be associated with the generator of the driven process. For convex LDF, equations (61) and (62) have the same solutions, i.e. if I is convex at $\boldsymbol{a}$, $(\boldsymbol{\omega_a},\boldsymbol{p_a})=(\boldsymbol{\omega_\gamma},\boldsymbol{p_\gamma})$ for $\boldsymbol{\gamma}=\boldsymbol{\nabla}I(a)$. This is in agreement with the equivalence of the microcanonical process and the conditioned-free process/driven process for $\boldsymbol{\gamma}=\boldsymbol{\nabla}I(a)$ [36]. In the following, we recover this result through direct calculation of the optimum of

$$\begin{aligned}\mathcal{F}(\boldsymbol{\omega},\boldsymbol{p})=&-I_{2.5}(\boldsymbol{\omega},\boldsymbol{p})+\gamma_1\,A_1(\boldsymbol{\omega})+\gamma_2\,A_2(\boldsymbol{p})\\&-\frac{1}{T}\int_0^T\mathrm{d}\tau\;c(\tau)\;\left[\;\sum_y p_y(\tau)-1\;\right]\\&-\frac{1}{T}\int_0^T\mathrm{d}\tau\sum_x u_x(\tau)\;\left[\;\dot{p}_x(\tau)-\sum_y\left(\omega_{xy}(\tau)-\omega_{yx}(\tau)\right)\;\right],\end{aligned} \tag{63}$$

where c and $\boldsymbol{u}$ are time dependent Lagrange multipliers respectively associated with the constraints **C1** and **C2**. We assume in addition that $\boldsymbol{u}(T)=\boldsymbol{u}(0)$. Functional derivatives with respect to occupation and transition probabilities yield

J. Stat. Mech. (2020) 033208

$$\begin{array}{lll} \frac{\partial \mathcal{F}}{\partial \omega_{xy}(\tau)} = 0 & \Rightarrow & \ln \frac{\omega_{xy}(\tau)}{k_{xy}(\tau)\, p_y(\tau)} + (u_y(\tau) - u_\tau) - \gamma_1\, g_{xy}(\tau) = 0 \qquad \text{for } x \neq y, \\ \frac{\partial \mathcal{F}}{\partial p_y(\tau)} = 0 & \Rightarrow & \sum_{x\neq y} \left[k_{xy}(\tau) - \frac{\omega_{xy}(\tau)}{p_y(\tau)} \right] + c(\tau) - \gamma_2\, f_y(\tau) \; - \dot{u}_y(\tau) = 0. \end{array} \tag{64}$$

We transform the first equation of (64) into

$$\omega_{xy}(\tau) = K'_{xy}(\tau)\; p_y(\tau), \tag{65}$$

with

$$K'_{xy}(\tau) \equiv k_{xy}(\tau)\; \mathrm{e}^{\gamma_1 g_{xy}(\tau)}\; \mathrm{e}^{u_x(\tau)-u_y(\tau)} = \kappa_{xy}(\tau)\mathrm{e}^{u_x(\tau)-u_y(\tau)}, \tag{66}$$

for $x \neq y$ and $\tau \in [0, T[$. We define the diagonal elements such that the sum over the lines of any column of $\boldsymbol{K'}(\tau)$ vanishes so that $\boldsymbol{K'}$ satisfies $\dot{\boldsymbol{p}} = \boldsymbol{K'}\boldsymbol{p}$ via condition **C2**:

$$K'_{yy}(\tau) \equiv -\sum_{x\neq y} K'_{xy}(\tau) \equiv \; - \; \Lambda'_y(\tau), \tag{67}$$

with $\Lambda'_y(\tau)$ the escape rate from state y at time τ associated with $\boldsymbol{K'}$. From condition **C3**, $\boldsymbol{p}$ is the TiPS probability associated with $\boldsymbol{K'}$. At this point, it remains to fix the Lagrange multipliers c and $\boldsymbol{u}$. As suggested by the notation, $\boldsymbol{K'}$ will turn out to be the generator $\boldsymbol{K}$ of the driven process defined in (39). To prove this, we see that the second equation of equation (64) becomes

$$c(\tau) = \sum_{x\neq y} K'_{xy} - \lambda_y(\tau) + \gamma_2\, f_y(\tau) + \dot{u}_y(\tau), \qquad \forall y. \tag{68}$$

Using (11) and (66), we get

$$c = \sum_{x\neq y} \kappa_{xy}\mathrm{e}^{u_x-u_y} + \kappa_{xx} + \dot{u}_y. \tag{69}$$

Multiplying by e^{u_y}, we finally obtain

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{\boldsymbol{u}}\right) = -\left(\mathrm{e}^{\boldsymbol{u}}\right)\left(\boldsymbol{\kappa} - c\mathbb{1}\right), \\ \mathrm{e}^{\boldsymbol{u}(0)} = \mathrm{e}^{\boldsymbol{u}(T)}, \end{cases} \tag{70}$$

with $(\mathrm{e}^{\boldsymbol{u}})_x \equiv \mathrm{e}^{u_x}$. The formal solution of (70) writes:

$$\mathrm{e}^{\boldsymbol{u}(t)} = \mathrm{e}^{\boldsymbol{u}(T)}\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}-c\mathbb{1}}(T,t) = \mathrm{e}^{\boldsymbol{u}(0)}\mathrm{e}^{-\int_t^T c}\; \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,t), \tag{71}$$

where we used property D.6 of appendix D in the second equality. Taking $t = 0$:

$$\mathrm{e}^{\boldsymbol{u}(0)}\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0) = \mathrm{e}^{\int_0^T c}\; \mathrm{e}^{\boldsymbol{u}(0)}. \tag{72}$$

Hence, the optimization with respect to $\boldsymbol{p}$ (second equation of (64)) leads to a spectral equation. Since the vector $\mathrm{e}^{\boldsymbol{u}(0)}$ has positive components, by Perron–Frobenius theorem it is the unique—up to a normalization—left eigenvector of $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0)$ associated with its largest eigenvalue ρ_T, hence

$$\mathrm{e}^{\int_0^T c} = \rho_T. \tag{73}$$

J. Stat. Mech. (2020) 033208

From equation (21), we find that the SCGF writes $\phi = \frac{1}{T}\int_0^T c$. We recover this result in appendix B directly from equation (62). Notice that we can rewrite equation (70) as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{\boldsymbol{u}+\int_t^T c}\right) = -(\mathrm{e}^{\boldsymbol{u}+\int_t^T c})\boldsymbol{\kappa}. \tag{74}$$

Hence the vector $\mathrm{e}^{\boldsymbol{u}+\int_t^T c}$ is solution of

$$\begin{cases} \dot{\boldsymbol{X}} = -\boldsymbol{X}\boldsymbol{\kappa}, \\ \boldsymbol{X}(T) = \rho_T^{-1}\boldsymbol{X}(0). \end{cases} \tag{75}$$

From equation (37), we conclude that the vector $\boldsymbol{l}$ that appears in the Doob transform leading to the driven generator is related to the Lagrange multipliers through

$$\boldsymbol{l}(t) = \mathrm{e}^{\boldsymbol{u}(t)+\int_t^T c}. \tag{76}$$

We emphasize that $\boldsymbol{u}(t)$ is set up to an additive and time-dependent function constant in the state space. Indeed, if **C2** is satisfied for all states but one then it is satisfied for all states (in view of **C1**). Then, equation (76) is a choice for this remaining degree of freedom in $\boldsymbol{u}(t)$.

We now show that the transition rate matrix $\boldsymbol{K'}$ generates the driven process, i.e. we show that $\boldsymbol{K'}$ writes as the Doob transform of $\boldsymbol{\kappa}$ associated with the vector $\boldsymbol{l} = \boldsymbol{l}(\tau)$. Using equations (11), (68) and (76), we transform equation (66) into

$$K'_{xy} = \kappa_{xy}\mathrm{e}^{u_x-u_y} - [\kappa_{xx} + \Lambda'_x]\,\delta_{xy}, \tag{77}$$

$$= \kappa_{xy}\mathrm{e}^{u_x-u_y} - [\,(-\lambda_x + \gamma_2 f_x) + (c + \lambda_x - \gamma_2\, f_x - \dot{u}_x)\,]\,\delta_{xy}, \tag{78}$$

$$= \mathrm{e}^{u_x}\kappa_{xy}\mathrm{e}^{-u_y} - \;[c - \dot{u}_x]\,\delta_{xy}, \tag{79}$$

$$= l_x\kappa_{xy}l_x^{-1} + l_x^{-1}\dot{l}_x\,\delta_{xy}, \tag{80}$$

$$= l_x\kappa_{xy}l_x^{-1} - l_x^{-1}(\boldsymbol{l}\boldsymbol{\kappa})_x\,\delta_{xy}, \tag{81}$$

$$= \kappa^{\boldsymbol{l}}_{xy}. \tag{82}$$

Hence, we conclude that $\boldsymbol{K'} = \boldsymbol{K}$ (defined in (39)). Then, the optimum of equation (62) is reached for $\boldsymbol{p} = \boldsymbol{\mu}$ the TiPS probability of the driven process with generator $\boldsymbol{K}$, and $\boldsymbol{\omega} = \boldsymbol{K} \circ \boldsymbol{p}$ the directional probability current associated with the probability $\boldsymbol{p}$ and rate matrix $\boldsymbol{K}$: in the time-homogeneous case, we recover the results of [33, 36].

To conclude this section, the driven process is the most probable process that reproduces the dynamics satisfying the imposed value of the conditioning observable. In other words, it is the generator of an 'optimal' Markov process for which the conditioning observable takes asymptotically the imposed value as a typical value.

J. Stat. Mech. (2020) 033208

5. Illustration on a solvable modulated two-level system

In this section, we consider a two level system with states denoted by $\mathbf{e}_\pm$. For simplicity, the transition rate matrix is chosen symmetric and piecewise-constant. We take as conditioning observable a current defined through a time-periodic function. We compute the SCGF and the rate matrices $\boldsymbol{\mathcal{K}}^n$ and $\boldsymbol{K}$ for the canonical and driven processes. We study the convergence of the canonical transition rates toward the driven one as the number of periods n grows. We also comment qualitatively on the influence of the conditioning on the transition rates of the driven process.

The transition rate matrix used to model the system writes

$$\boldsymbol{k}(t) = \begin{pmatrix} -k(t) & k(t) \\ k(t) & -k(t) \end{pmatrix}. \tag{83}$$

The rate $k(t)$ is a T-periodic and piecewise constant function of time

$$k(t) = \begin{cases} k^0 & \text{for } t \in [0, \alpha T[, \\ k^1 & \text{for } t \in [\alpha T, T[, \end{cases} \tag{84}$$

where $k^i > 0$, $i = 0, 1$ are two constants. We chose $k^0 = 1$ to set the time scale. We use $\alpha \in]0, 1[$ as duty cycle of the piecewise modulation. The observable A is the scalar path functional

$$A[z] = \frac{1}{nT} \sum_{t \in [0, nT] \,|\, z(t^+) \neq z(t^-)} g_{z(t^+), z(t^-)}(t), \tag{85}$$

where we assume that $\boldsymbol{g}$ is antisymmetric and piecewise constant with duty cycle α, i.e. $g_{-+}(t) = -g_{+-}(t) \equiv g(t)$ and

$$g(t) = \begin{cases} g^0 & \text{for } t \in [0, \alpha T[, \\ g^1 & \text{for } t \in [\alpha T, T[. \end{cases} \tag{86}$$

When $g^0 = 1$ and $g^1 = 0$ for instance, A counts the net number of transitions $\mathbf{e}_+ \to \mathbf{e}_-$ occuring in the first part of each period. With $\mathbf{e}_+$, $\mathbf{e}_-$ respectively the first and second basis vectors, the tilted operator associated to $\boldsymbol{k}$ and observable A writes

$$\boldsymbol{\kappa}(\boldsymbol{\gamma}, t) = \begin{pmatrix} -k(t) & k(t)\mathrm{e}^{-\gamma g(t)} \\ k(t)\mathrm{e}^{\gamma g(t)} & -k(t) \end{pmatrix}. \tag{87}$$

Our theory relies on the propagator $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t, 0)$ that we shall now obtain to proceed. The tilted operator being piecewise constant, this propagator can be writen explicitly. For $t \in [0, \alpha T[$, we obtain

$$\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t, 0) = \mathrm{e}^{-k^0 t} \begin{pmatrix} \cosh(k^0 t) & \mathrm{e}^{-\gamma g^0} \sinh(k^0 t) \\ \mathrm{e}^{\gamma g^0} \sinh(k^0 t) & \cosh(k^0 t) \end{pmatrix}, \tag{88}$$

while for $t \in [\alpha T, T[$, and introducing $t^0 \equiv \alpha T$ and $t^1 \equiv t^1(t) \equiv t - \alpha T$, we have

J. Stat. Mech. (2020) 033208

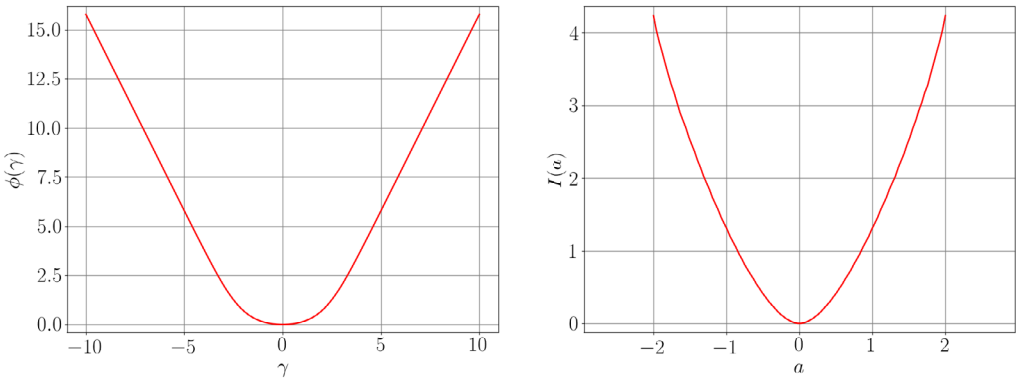

Figure 1. (left) SCGF $\phi(\gamma)$ and (right) LDF $I(a)$. The figures are obtained for $\alpha = 0.3,\ T = 1,\ k^0 = 1,\ k^1 = 0.1,\ g^0 = 1,\ g^1 = -1$.

$$\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,0) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t^0+t^1,0) = \mathrm{e}^{-k^0t^0-k^1t^1} \times \begin{pmatrix} \prod_i \cosh(k^it^i) + \prod_i \mathrm{e}^{\gamma(1-2i)g^i}\sinh(k^it^i) & \sum_i \mathrm{e}^{-\gamma g^i}\sinh(k^it^i)\cosh(k^{1-i}t^{1-i}) \\ \sum_i \mathrm{e}^{\gamma g^i}\sinh(k^it^i)\cosh(k^{1-i}t^{1-i}) & \prod_i \cosh(k^it^i) + \prod_i \mathrm{e}^{-\gamma(1-2i)g^i}\sinh(k^it^i) \end{pmatrix}, \tag{89}$$

where sums and products are on $i = 0, 1$.

The largest eigenvalue of the propagator over one period writes

$$\rho_T = \frac{1}{2}\left[\mathrm{tr}\ \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0) + \sqrt{\left[\mathrm{tr}\ \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)\right]^2 - 4\det\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)}\right], \tag{90}$$

where tr and det stands for the trace and determinant, respectively. Using equation (21), the SCGF $\phi(\gamma)$ follows, see figure 1 for a numerical computation. The Legendre conjugate LDF $I(a)$ is shown on the same figure. Notice that I vanishes at $a = 0$ due to the symmetry of the rate matrix $\boldsymbol{k}$: there are asymptotically as many transitions $\mathbf{e}_+ \to \mathbf{e}_-$ than transitions $\mathbf{e}_- \to \mathbf{e}_+$ leading to a vanishing typical value for A.

The generator $\boldsymbol{K}$ being defined as the Doob transform of $\boldsymbol{\kappa}$ based on $\boldsymbol{l}(t) = \boldsymbol{l}(0)[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,0)]^{-1}$, we need the left eigenvector $\boldsymbol{l}(0) = \boldsymbol{l}_T$ of the one period propagator associated to the eigenvalue ρ_T:

$$\boldsymbol{l}(0) = \frac{1}{\mathcal{N}}\begin{pmatrix} \prod_i \mathrm{e}^{-k^it^i}\left[\sum_i \mathrm{e}^{\gamma g^i}\sinh(k^it^i)\cosh(k^{1-i}t^{1-i})\right] \\ \rho_T - \prod_i \mathrm{e}^{-k^it^i}\left[\prod_i \cosh(k^it^i) + \prod_i \mathrm{e}^{\gamma(1-2i)g^i}\sinh(k^it^i)\right] \end{pmatrix}^{\dagger}, \tag{91}$$

with $\mathcal{N}$ a normalizing factor following from equation (17). Inverting the propagators in equations (88) and (89), we can compute $\boldsymbol{l}(t)$ at any $t \in [0, T[$. Then, equation (39) yields an analytic expression for the generator of the driven process from which we have computed numerically one component shown on figure 2. Similarly, the generator $\boldsymbol{\mathcal{K}}^n$ is defined as the Doob transform of $\boldsymbol{\kappa}$ based on $\boldsymbol{C}^n(t) = \mathbf{1}\ [\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)]^n[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,0)]^{-1}$. Inverting the propagators of equations (88) and (89) and taking the nth power of the one period propagator, we can compute $\boldsymbol{C}^n(t)$ at any $t \in [0, T[$ for $n \in \mathbb{N}$. Then, equation (28) yields an expression for the generator of the canonical process from which

J. Stat. Mech. (2020) 033208

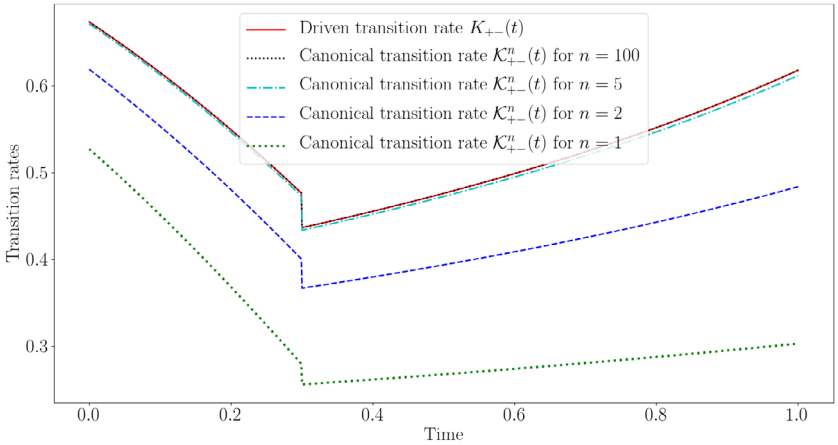

Figure 2. $K_{+-}(t)$ and $\mathcal{K}^n_{+-}(t)$ as a function of time for different number of periods $n = 1, 2, 5, 100$. The figure is obtained for $\alpha = 0.3$, $T = 1$, $k^0 = 1$, $k^1 = 0.1$, $g^0 = 1$, $g^1 = -1$, $a = 0.4$ corresponding to $\gamma = 1.11$.

we have computed numerically one component shown on figure 2 for $n = 1, 2, 5$ and 100. This figure illustrates the convergence of the canonical generator $\boldsymbol{\mathcal{K}}^n$ towards the driven generator $\boldsymbol{K}$ when $n \to \infty$ as stated in equation (40). We observe that the two generators $\boldsymbol{K}$ and $\boldsymbol{\mathcal{K}}^n$ are piecewise continuous (with discontinuities at phases αT and T) and time-dependent even though the original rate matrix $\boldsymbol{k}$ was piecewise constant.

On figure 3, we plot both driven rates $K_{+-}(t)$ and $K_{-+}(t)$ and original rates $k_{+-}(t) = k_{-+}(t) = k(t)$ to observe qualitatively the effect of the conditioning on our initial Markov process. We chose to impose $a = 0.4$ net transitions from $\mathbf{e}_+$ to $\mathbf{e}_-$ per unit time, counted positively if they occur on the first part of each cycle ($g^0 = 1$) and negatively on the second part ($g^1 = -1$). In view of I's strict convexity, the process that has $a = 0.4$ as a typical event is the driven process for $\gamma = 1.11 = I'(0.4)$. In the original process, A is zero on average due to the symmetry of the rate matrix. Hence, imposing $a > 0$ should increase the rate of the driven process for transitions $\mathbf{e}_+ \to \mathbf{e}_-$ on $[0, \alpha T[$ and transitions $\mathbf{e}_- \to \mathbf{e}_+$ on $[\alpha T, T]$. Compared to the original rate k, we see on figure 3 that indeed $K_{+-} < k^0 < K_{-+}$ on $[0, \alpha T[$ so that transitions $\mathbf{e}_+ \to \mathbf{e}_-$ are prefered on average, and conversely $K_{-+} < k^1 < K_{+-}$ on $[\alpha T, T]$ so that transitions $\mathbf{e}_- \to \mathbf{e}_+$ are prefered on average. Hence, the conditioning has broken the symmetry of the rate matrix and made it fully time dependent.

On figure 4, we plot the rate K_{+-} for different values of a (associated to their corresponding γ). We observe that this rate from $\mathbf{e}_- \to \mathbf{e}_+$ deviates more and more from k_{+-} as $|a|$ becomes larger, i.e. goes away from the mean 0 for the original process. The magnitude of change of the driven rate is thus in direct correspondence with the magnitude of the conditioning. However, it is not intuitive to understand the growth of the transition rate. We can just say that the possibility of a time-dependent rate matrix offers a broader dynamical space to explore in the variational calculation compared to case of piecewise-constant rates.

J. Stat. Mech. (2020) 033208

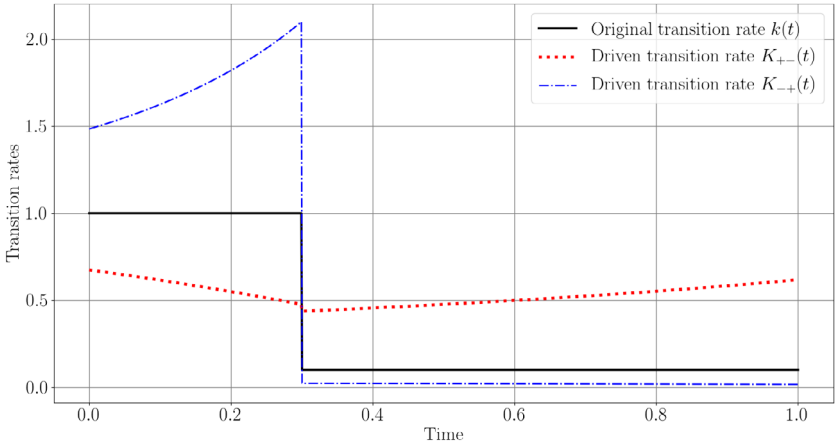

Figure 3. Original transition rate $k(t)$ (solid black line) and driven transition rates $K_{+-}(t)$ (red dotted line) and $\boldsymbol{K}_{-+}(t)$ (blue dashed line). The figure is obtained for $\alpha = 0.3,\ T = 1,\ k^0 = 1,\ k^1 = 0.1,\ g^0 = 1,\ g^1 = -1,\ a = 0.4$ corresponding to $\gamma = 1.11$.

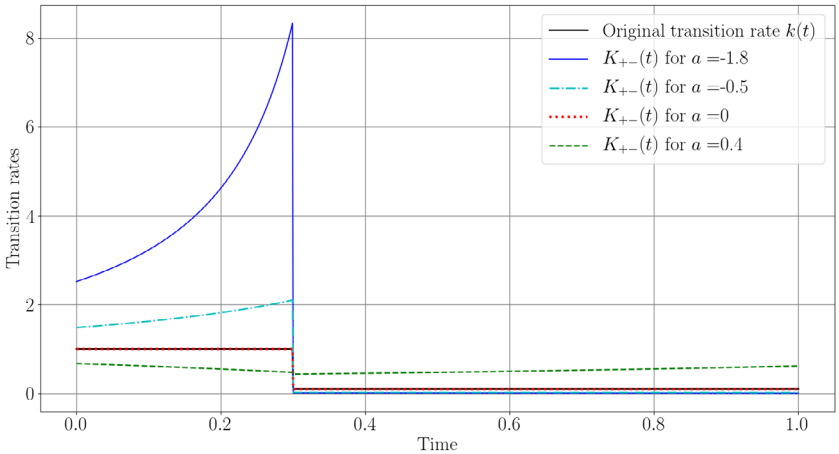

Figure 4. Original transition rate $k(t)$ (solid black line) and driven rate $K_{+-}(t)$ (colored lines) for different values of $\boldsymbol{\gamma}$ corresponding to different values of a. The figure is obtained for $\alpha = 0.3,\ T = 1,\ k^0 = 1,\ k^1 = 0.1,\ g^0 = 1,\ g^1 = -1$. The values $a = -1.8, -0.5, 0, 0.4$ correspond to $\gamma = -3.59, -1.29, 0, 1.11$, respectively. As expected, $\boldsymbol{k} = \boldsymbol{K}$ when conditioning at the mean value $a = 0$.

6. Conclusion

Beyond the computation of the cumulants of a random variable $\boldsymbol{A}$, recent developments in large deviation theory provide a mathematical framework to study path probabilities conditioned on an event $\boldsymbol{A} = \boldsymbol{a}$ (microcanonical conditioning, $\boldsymbol{A}$ does not fluctuate). One aims to build a new Markov process for which $\boldsymbol{A}$ converges in probability to the value $\boldsymbol{a}$ at long time (canonical conditioning, $\boldsymbol{A}$ fluctuates). In this paper, we addressed this problem of process conditioning for observables defined through periodic functions in the framework of Markov jump processes with time-periodic generators. We took the period of these functions equal to the period of the generator, with no loss of generality compared to the case of commensurable periods. We focused on jump processes,

J. Stat. Mech. (2020) 033208

but we expect our results to be transposable to general Markov processes. Starting from nonequilibrium path probabilities generalizing the canonical and microcanonical ensembles, we defined the Markov generator of the canonical process and its asymptotic equivalent after a large number of periods. The latter is the driven generator obtained from the Doob transform involving an eigenvector of the one-period propagator for the tilted operator (and its time evolution). This is consistent with the time-homogeneous theory where an eigenvector of the tilted matrix is involved instead. Finally, the conditioned-free process for which $\boldsymbol{A}$ takes asymptotically the microcanonical value $\boldsymbol{a}$ follows from the driven process. This result requires the ensemble equivalence between microcanonical and canonical path ensembles which is guaranteed by the convexity of the LDF for $\boldsymbol{A}$, in straight connection with entropy's concavity for the equivalence of equilibrium ensembles. This analogy between entropy and LDF is broader than the question of ensemble equivalence. In the same way that the canonical state probability follows from Jayne's Maximum entropy principle in equilibrium statistical mechanics, the driven process follows from a constrained optimization problem on the 2.5 LDF.

Acknowledgment

We thank H Vroylandt for his pertinent comments on the definition of the Doob transform via the variational approach. The work of R Chetrite has been supported by the project LSD ANR-15-CE40-0020-01 of the French National Research Agency (ANR).

Appendix A. Definition of the Doob transform

Let $\boldsymbol{M}$ be an arbitrary Metzler matrix [62] with nonnegative off-diagonal components and $\boldsymbol{v}$ a vector whose elements are strictly positive. The Doob transform of $\boldsymbol{M}$ associated with $\boldsymbol{v}$ is defined by

$$\boldsymbol{M}^{\boldsymbol{v}} \equiv \mathcal{D}(\boldsymbol{v})\boldsymbol{M}\,\mathcal{D}(\boldsymbol{v})^{-1} - \mathcal{D}(\boldsymbol{v})^{-1}\mathcal{D}(\boldsymbol{v}\boldsymbol{M}), \tag{A.1}$$

where $\mathcal{D}(\boldsymbol{v})$ is the diagonal matrix with the components of $\boldsymbol{v}$ on its diagonal. Componentwise, equation (A.1) writes

$$M^{\boldsymbol{v}}_{xy}(t) = v_x(t)M_{xy}(t)v_y^{-1}(t) - v_x^{-1}(t)(\boldsymbol{v}\boldsymbol{M})_x(t)\delta_{xy}. \tag{A.2}$$

$\boldsymbol{M}^{\boldsymbol{v}}$ is a norm conserving generator since $\sum_x M^{\boldsymbol{v}}_{xy} = 0$ for all y. The Doob transform is then a tool to build a norm conserving Markov generator out of a Metzler matrix. The above definition of the Doob transform is a special case of a more general definition, see [35]. We emphasize that the norm of the vector used in a Doob transform plays no role: For $\alpha \equiv \alpha(t)$ a scalar function of time, the Doob transforms based on $\boldsymbol{v}(t)$ and $\alpha(t)\boldsymbol{v}(t)$ are the same, i.e. $\boldsymbol{M}^{\alpha\boldsymbol{v}} = \boldsymbol{M}^{\boldsymbol{v}}$. The path probability associated with the Doob transform $\boldsymbol{M}^{\boldsymbol{v}}$ is given by [63]

$$\begin{aligned}\mathbb{P}_{\boldsymbol{M}^{\boldsymbol{v}},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{M},\boldsymbol{\pi}(0)}[z]\, v_{x(nT)}(nT)\\ \exp\left[-\int_0^{nT}\left(v^{-1}_{x(t)}(t)(\boldsymbol{v}\boldsymbol{M})_{x(t)}(t) + v^{-1}_{x(t)}\left.\frac{\partial \boldsymbol{v}(t)}{\partial t}\right|_{x(t)}\right)\mathrm{d}t\right] v^{-1}_{x_0}(0),\end{aligned} \tag{A.3}$$

J. Stat. Mech. (2020) 033208

where the second term in the integrand is due to the time dependence of vector $\boldsymbol{v}$ leading to $\ln v_{x_i}(t_{i+1}) - \ln v_{x_i}(t_i) = \int_{t_i}^{t_{i+1}} dt \partial_t \ln v_{x_i}(t)$ contributions for each interval of time $[t_i, t_{i+1}]$ between two jumps.

Appendix B. SCGF from the optimizer

We recover the SCGF for observable $\boldsymbol{A}$ by evaluating the 2.5 LDF at the optimum $(\boldsymbol{\omega}, \boldsymbol{p})$ of our variational problem stated at equation (61). By definition of the second component of our observable $\boldsymbol{A}$ in equation (58) and using equation (68) we find

$$\gamma_2\ A_2(\boldsymbol{p}) = \frac{1}{T}\int_0^T \sum_x p_x\ \gamma_2\, f_x, \tag{B.1}$$

$$= \frac{1}{T}\int_0^T \left[c + \sum_x p_x\left[\lambda_x - \Lambda_x\right] - \sum_x p_x \dot{u}_x\right], \tag{B.2}$$

$$= \frac{1}{T}\int_0^T \left[c + \sum_x p_x\left[\lambda_x - \Lambda_x\right] + \sum_x \dot{p}_x u_x\right], \tag{B.3}$$

where we used **C3** in the integration by part. Using equations (65), (66) and **C2**, the LDF at the optimum writes

$$I(\boldsymbol{\omega}, \boldsymbol{p}) = \frac{1}{T}\int_0^T \left[\sum_x p_x\left[\lambda_x - \Lambda_x\right] + \sum_{x, y\neq x} \omega_{xy}\left[u_x - u_y\right] + \gamma_1 \sum_{x, y\neq x} \omega_{xy}\, g_{xy}\right], \tag{B.4}$$

$$= \frac{1}{T}\int_0^T \left[\sum_x p_x\left[\lambda_x - \Lambda_x\right] + \sum_{x, y\neq x} u_x\left[\omega_{xy} - \omega_{yx}\right] + \gamma_1 \sum_{x, y\neq y} \omega_{xy}\, g_{xy}\right], \tag{B.5}$$

$$= \frac{1}{T}\int_0^T \left[\sum_x p_x\left[\lambda_x - \Lambda_x\right] + \sum_x u_x\, \dot{p}_x\right] + \gamma_1 A_1. \tag{B.6}$$

Combining equations (B.3) and (B.6), we finally obtain

$$\gamma_1\ A_1(\boldsymbol{K}\circ\boldsymbol{p}) + \gamma_2\ A_2(\boldsymbol{p}) - I_{2.5}(\boldsymbol{K}\circ\boldsymbol{p}, \boldsymbol{p}) = \frac{1}{T}\int_0^T c(\tau)\mathrm{d}\tau = \phi(\boldsymbol{\gamma}). \tag{B.7}$$

The left-hand-side is the SCGF as the Legendre transform of the LDF. It follows that the SCGF is the time-average over a period of the Lagrange multiplier used to normalize the occupation density, recovering the result stated in equation (73). As mentioned in the conclusion, the variational calculation of the SCGF is similar in many ways to the calculation of equilibrium canonical probability via the maximum entropy principle in which the SCGF (free energy) is also connected to the Lagrange multiplier that imposes probability normalization.

J. Stat. Mech. (2020) 033208

Appendix C. Definitions of the time-ordered exponential

The ordered exponential $\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0) \equiv \overleftarrow{\exp} \int_0^t \boldsymbol{M}(t')\,\mathrm{d}t'$ is the unique solution of the initial matrix value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{M}(t)\boldsymbol{X}(t), \quad \text{with } \boldsymbol{X}(0) = \mathbb{1}, \tag{C.1}$$

that has the integral form

$$\boldsymbol{X}(t) = \mathbb{1} + \int_0^t \boldsymbol{M}(t')\boldsymbol{X}(t')\mathrm{d}t'. \tag{C.2}$$

Inserting this integral form into itself, one obtains the series expansion of the time ordered exponential

$$\begin{aligned}\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(\tau,0) = \mathbb{1} &+ \int_0^t \boldsymbol{M}(\tau_1)\mathrm{d}\tau_1 + \int_0^t \mathrm{d}\tau_1 \int_0^{t_1} \mathrm{d}\tau_2\, \boldsymbol{M}(\tau_1)\boldsymbol{M}(\tau_2) \\ &+ \int_0^t \mathrm{d}\tau_1 \int_0^{t_1} \mathrm{d}\tau_2 \int_0^{t_2} \mathrm{d}\tau_3\, \boldsymbol{M}(\tau_1)\boldsymbol{M}(\tau_2)\boldsymbol{M}(\tau_3) + \dots .\end{aligned} \tag{C.3}$$

Notice that the arrow on the exponential specifies the ordering of the product of $\boldsymbol{M}$ in the expansion for increasing time from right to left (having in mind that $\tau_n < \dots < \tau_3 < \tau_2 < \tau_1$).

The reverse-ordered exponential $\overrightarrow{\mathcal{Q}}_{\boldsymbol{M}}(0,t) \equiv \overrightarrow{\exp} \int_0^t \boldsymbol{M}(t')\,\mathrm{d}t'$ is unique solution of the initial matrix value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{X}(t)\boldsymbol{M}(t), \quad \text{with } \boldsymbol{X}(0) = \mathbb{1}, \tag{C.4}$$

that has the integral form

$$\boldsymbol{X}(t) = \mathbb{1} + \int_0^t \boldsymbol{X}(t')\boldsymbol{M}(t')\mathrm{d}t'. \tag{C.5}$$

Inserting this integral form into itself, one obtains the series expansion of the reverse-ordered exponential

$$\begin{aligned}\overrightarrow{\mathcal{Q}}_{\boldsymbol{M}}(0,\tau) = \mathbb{1} &+ \int_0^t \boldsymbol{M}(\tau_1)\mathrm{d}\tau_1 + \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2\, \boldsymbol{M}(\tau_2)\boldsymbol{M}(\tau_1) \\ &+ \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \int_0^{\tau_2} \mathrm{d}\tau_3\, \boldsymbol{M}(\tau_3)\boldsymbol{M}(\tau_2)\boldsymbol{M}(\tau_1) + \dots .\end{aligned} \tag{C.6}$$

Notice that the arrow on the exponential specifies the ordering of the product of $\boldsymbol{M}$ in the expansion for increasing time from left to right (we recall that $\tau_n < \dots < \tau_3 < \tau_2 < \tau_1$).

Appendix D. Properties of the time-ordered exponential

For the reader convenience, we recall useful properties on linear differential equations with periodic generators. See [64] for a full description of the theory.

J. Stat. Mech. (2020) 033208

Property D.1 (Transpose of a propagator). The transpose of a propagator based on generator $\boldsymbol{M}$ is the time reverse propagator based on the transposed generator $\boldsymbol{M}^\dagger$

$$\left[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(t,0)\right]^\dagger = \overrightarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}^\dagger}(0,t). \tag{D.1}$$

The property follows from the definitions and the fact that the transpose of a product of two matrices is the product of the two transposed matrices taken in the reverse order.

Property D.2 (Inverse of a propagator). The inverse of a propagator based on generator $\boldsymbol{M}$ is the time reverse propagator based on the opposite generator $-\boldsymbol{M}$

$$\left[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(t,0)\right]^{-1} = \overrightarrow{\boldsymbol{\mathcal{Q}}}_{-\boldsymbol{M}}(0,t). \tag{D.2}$$

Assuming $\dot{\boldsymbol{X}} = \boldsymbol{M}\boldsymbol{X}$ and since $\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{X}\boldsymbol{X}^{-1}) = \dot{\boldsymbol{X}}\boldsymbol{X}^{-1} + \boldsymbol{X}\dot{\boldsymbol{X}}^{-1} = 0$, we have $\dot{\boldsymbol{X}}^{-1} = -\boldsymbol{X}^{-1}\dot{\boldsymbol{X}}\boldsymbol{X}^{-1} = -\boldsymbol{X}^{-1}\boldsymbol{M}$. Hence the two propagators are connected.

Property D.3 (First relation between final and initial value problems). The solution of the final value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = -\boldsymbol{X}(t)\boldsymbol{M}(t), \quad \text{with } \boldsymbol{X}(T) = \mathbb{1}, \tag{D.3}$$

is given by $\boldsymbol{X}(t) = \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t)$.

Indeed, one can check directly that

$$\frac{\mathrm{d}}{\mathrm{d}t}\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t) = \lim_{a\to 0}\frac{\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t+a) - \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t+a)\,\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(t+a,t)}{a}, \tag{D.4}$$

$$= \lim_{a\to 0}\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t+a) \times \frac{\mathbb{1} - \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(t+a,t)}{a}, \tag{D.5}$$

$$= \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t) \times \left[-\frac{\mathrm{d}}{\mathrm{d}s}\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(s,t)\,|_{s=t}\right], \tag{D.6}$$

$$= -\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t)\;\boldsymbol{M}\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(t,t), \tag{D.7}$$

$$= -\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t)\;\boldsymbol{M}. \tag{D.8}$$

Property D.4 (Second relation between final and initial value problems). The solution of the final matrix value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{M}(t)\boldsymbol{X}(t), \quad \text{with } \boldsymbol{X}(T) = \mathbb{1}, \tag{D.9}$$

is given by $\boldsymbol{X}(t) = \left[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{M}}(T,t)\right]^{-1} = \overrightarrow{\boldsymbol{\mathcal{Q}}}_{-\boldsymbol{M}}(t,T)$.

This follows from combining properties D.2 and D.3.

J. Stat. Mech. (2020) 033208

Property D.5 (Multiplicative property of propagators). For any $t_1 \in [t_0, t]$,

$$\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t, t_0) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t, t_1)\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_1, t_0). \tag{D.10}$$

Property D.6 (Time-ordered exponential of the sum of commuting matrices). If $\boldsymbol{M}(t_1)$ and $\boldsymbol{N}(t_2)$ commute for any t_1,$t_2 \in \mathbb{R}$, then $\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}+\boldsymbol{N}}(t, t_0) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t, t_0)\overleftarrow{\mathcal{Q}}_{\boldsymbol{N}}(t, t_0)$.

Let us denote the left-hand side of the equality by $\boldsymbol{X}(t)$ and the right-hand side by $\boldsymbol{Y}(t)$. On the one hand, $\dot{\boldsymbol{X}} = (\boldsymbol{M} + \boldsymbol{N})\boldsymbol{X}$. On the other hand, $\dot{\boldsymbol{Y}} = \boldsymbol{M}\boldsymbol{Y} + \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t, t_0)\boldsymbol{N}\overleftarrow{\mathcal{Q}}_{\boldsymbol{N}}(t, t_0) = (\boldsymbol{M} + \boldsymbol{N})\boldsymbol{Y}$ since $\boldsymbol{M}$ and $\boldsymbol{N}$ commute for any time. Thus, the matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ satisfy the same matrix differential equation. Besides, $\boldsymbol{X}(t_0) = \boldsymbol{Y}(t_0) = \mathbb{1}$, hence $\boldsymbol{X}(t) = \boldsymbol{Y}(t)$, $\forall t \in \mathbb{R}$.

References

[1] Dembo A and Zeitouni O 1998 *Large Deviations Techniques and Applications* (Berlin: Springer)
[2] Amann A and Atmanspacher H 1999 *J. Sci. Explor.* **13** 639–64
[3] den Hollander F 2000 *Fields Inst. Monogr.* 14 (www.ams.org/publications/authors/books/postpub/fim-14)
[4] Ellis R S 1985 *Entropy, Large Deviations, and Statistical Mechanics* (New York: Springer)
[5] Touchette H 2009 *Phys. Rep.* **478** 1–69
[6] Beck C and Schögl F 1993 *Thermodynamics of Chaotic Systems: an Introduction* vol 4 (Cambridge: Cambridge University Press)
[7] Gaspard P 1998 *Chaos, Scattering and Statistical Mechanics* (*Cambridge Nonlinear Science Series*) (Cambridge: Cambridge University Press)
[8] Callen H B 1985 *Thermodynamics and an Introduction to Thermostatistics* 2nd edn (New York: Wiley)
[9] Deuschel J D, Stroock D W and Zessin H 1991 *Commun. Math. Phys.* **139** 83–101
[10] Dobrushin R L, Kotecký R and Shlosman S 1992 *Wulff Construction: a Global Shape from Local Interaction* vol 104 (Providence, RI: American Mathematical Society)
[11] Donnelly P and Ellis R S 1987 *J. Am. Stat. Assoc.* **82** 948
[12] Lanford O E 1973 Entropy and equilibrium states in classical statistical mechanics *Statistical Mechanics and Mathematical Problems* (Berlin: Springer) pp 1–113
[13] Ruelle D 1999 *Statistical mechanics: Rigorous Results* (Singapore: World Scientific)
[14] Chetrite R and Touchette H 2013 *Phys. Rev. Lett.* **111** 120601
[15] Grassberger P 1989 *J. Phys. A: Math. Gen.* **22** 3283
[16] Hamm A, Tél T and Graham R 1994 *Phys. Lett.* A **185** 313–20
[17] Grafke T and Vanden-Eijnden E 2017 *J. Stat. Mech.* 093208
[18] Garrahan J P, Jack R L, Lecomte V, Pitard E, van Duijvendijk K and van Wijland F 2007 *Phys. Rev. Lett.* **98** 195702
[19] Chandler D and Garrahan J P 2010 *Annu. Rev. Phys. Chem.* **61** 191–217
[20] Jack R L and Sollich P 2013 *J. Phys. A: Math. Theor.* **47** 015003
[21] Lazarescu A, Cossetto T, Falasco G and Esposito M 2019 *J. Chem. Phys.* **151** 064117
[22] Dellago C, Bolhuis P G, Csajka F S and Chandler D 1998 *J. Chem. Phys.* **108** 1964–77
[23] Hedges L O, Jack R L, Garrahan J P and Chandler D 2009 *Science* **323** 1309–13
[24] Giardina C, Kurchan J, Lecomte V and Tailleur J 2011 *J. Stat. Phys.* **145** 787–811
[25] Nemoto T and Sasa S I 2014 *Phys. Rev. Lett.* **112** 090602
[26] Ferré G, Touchette H, Ferre G and Touchette H 2018 *J. Stat. Phys.* **172** 1525–44
[27] Grafke T and Vanden-Eijnden E 2019 *Chaos* **29** 063118
[28] Chauvin B and Rouault A 1988 *Probab. Theory Relat. Fields* **80** 299–314
[29] Derrida B and Shi Z 2016 *J. Stat. Phys.* **163** 1285–311
[30] Pham H 2007 Some applications and methods of large deviations in finance and insurance *Paris-Princeton Lectures on Mathematical Finance 2004* (Berlin: Springer) pp 191–244
[31] Arratia R, Gordon L and Waterman M S 1990 *Ann. Stat.* **18** 539–70
[32] Wachtel A, Vollmer J and Altaner B 2015 *Phys. Rev.* E **92** 042132
[33] Verley G 2016 *Phys. Rev.* E **93** 012111

J. Stat. Mech. (2020) 033208

[34] Jack R L and Sollich P 2010 *Prog. Theor. Phys. Suppl.* **184** 304–17
[35] Chetrite R and Touchette H 2015 *Ann. Henri Poincaré* **16** 2005–57
[36] Chetrite R and Touchette H 2015 *J. Stat. Mech.* P12001
[37] Blickle V and Bechinger C 2012 *Nat. Phys.* **8** 143
[38] Martínez I A, Roldán E, Dinis L, Petrov D, Parrondo J M and Rica R A 2016 *Nat. Phys.* **12** 67
[39] Erbas-Cakmak S, Leigh D A, McTernan C T and Nussbaumer A L 2015 *Chem. Rev.* **115** 10081–206
[40] Martínez I A, Roldán E, Dinis L and Rica R A 2017 *Soft Matter* **13** 22–36
[41] Verley G, Chétrite R and Lacoste D 2011 *J. Stat. Mech.* P10025
[42] Verley G, den Broeck C V and Esposito M 2014 *New J. Phys.* **16** 095001
[43] Barato A C and Chetrite R 2018 *J. Stat. Mech.* 053207
[44] Li Y 2016 *Nonlinear Anal.: Real World Appl.* **28** 32–47
[45] Andrieux D and Gaspard P 2008 *J. Chem. Phys.* **128** 154506
[46] Herpich T, Thingna J and Esposito M 2018 *Phys. Rev.* X **8** 031056
[47] Van Kampen N G 2007 *Stochastic Processes in Physics and Chemistry* (Oxford: Elsevier)
[48] Anderson W J 1991 *Continuous-Time Markov Chains: an Applications-Oriented Approach* (New York: Springer)
[49] Derrida B and Lebowitz J L 1998 *Phys. Rev. Lett.* **80** 209
[50] Lebowitz J L and Spohn H 1999 *J. Stat. Phys.* **95** 333–65
[51] Verley G, Van den Broeck C and Esposito M 2013 *Phys. Rev.* E **88** 032137
[52] Chetrite R 2018 *HDR Thesis* Université de Nice-Sophia-Antipolis, France
[53] Touchette H 2015 *J. Stat. Phys.* **159** 987–1016
[54] Garrahan J P, Jack R L, Lecomte V, Pitard E, van Duijvendijk K and van Wijland F 2009 *J. Phys. A: Math. Theor.* **42** 075007
[55] Ferrario M, Ciccotti G and Binder K 2006 *Computer Simulations in Condensed Matter: from Materials to Chemical Biology. Volume* 1 (New York: Springer)
[56] Bertini L, Chetrite R, Faggionato A and Gabrielli D 2018 *Ann. Henri Poincaré* **19** 3197–238
[57] Singh N and Wynants B 2010 *J. Stat. Mech.* P03007
[58] Maes C and Netočný K 2008 *Europhys. Lett.* **82** 30003
[59] Barato A C and Chétrite R 2015 *J. Stat. Phys.* **160** 1154–72
[60] Jack R L and Sollich P 2015 *Eur. Phys. J.-Spec. Top.* **224** 2351–67
[61] Hoppenau J, Nickelsen D and Engel A 2016 *New J. Phys.* **18** 083010
[62] Farina L and Rinaldi S 2000 *Positive Linear Systems: Theory and Applications* (New York: Wiley)
[63] Chetrite R and Gupta S 2011 *J. Stat. Phys.* **143** 543–84
[64] Adrianova L Y 1995 *Introduction to Linear Systems of Differential Equation* (Providence, RI: American Mathematical Society)

J. Stat. Mech. (2020) 033208