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

Joint probability distributions and fluctuation theorems

Reinaldo García-García[1,3], **Vivien Lecomte**[2], **Alejandro B Kolton**[1] **and Daniel Domínguez**[1]

[1] Centro Atómico Bariloche and Instituto Balseiro, 8400 S C de Bariloche, Argentina
[2] Laboratoire de Probabilités et Modèles Aléatoires (CNRS UMR 7599), Université Pierre et Marie Curie—Paris VI and Université Paris-Diderot—Paris VII, site Chevaleret, 175 rue du Chevaleret, F-75013 Paris, France
E-mail: reinaldo.garcia@cab.cnea.gov.ar, vivien.lecomte@univ-paris-diderot.fr, koltona@cab.cnea.gov.ar and domingd@cab.cnea.gov.ar

Received 28 November 2011
Accepted 16 January 2012
Published 9 February 2012

Online at stacks.iop.org/JSTAT/2012/P02009
doi:10.1088/1742-5468/2012/02/P02009

Abstract. We derive various exact results for Markovian systems that spontaneously relax to a non-equilibrium steady state by using joint probability distribution symmetries of different entropy production decompositions. The analytical approach is applied to diverse problems such as the description of the fluctuations induced by experimental errors, for unveiling symmetries of correlation functions appearing in fluctuation–dissipation relations recently generalized to non-equilibrium steady states, and also for mapping averages between different trajectory-based dynamical ensembles. Many known fluctuation theorems arise as special instances of our approach for particular twofold decompositions of the total entropy production. As a complement, we also briefly review and synthesize the variety of fluctuation theorems applying to stochastic dynamics of both continuous systems described by a Langevin dynamics and discrete systems obeying a Markov dynamics, emphasizing how these results emerge from distinct symmetries of the dynamical entropy of the trajectory followed by the system. For Langevin dynamics, we embed the 'dual dynamics' with a physical meaning, and for Markov systems we show how the fluctuation theorems translate into symmetries of modified evolution operators.

Keywords: stochastic particle dynamics (theory), fluctuations (theory), stationary states

[3] Author to whom any correspondence should be addressed.

Contents

1. Introduction

The study of the fluctuating heat interchange between a small system and a thermal reservoir is of academic interest but also of direct experimental relevance, as new techniques for microscopic manipulation and detection allow us nowadays to measure fluctuations in small experimental systems of relevance in physics, chemistry and biology [1, 2]. In this respect, a group of relations known as fluctuation theorems (FT) [3]–[11] have attracted a lot of attention as they shed new light into the principles governing energy fluctuations in a family of model systems. Remarkably, these results go beyond linear-response or quasi-equilibrium conditions, and apply to systems driven by non-conservative forces with arbitrary time-dependent protocols, even feedback-controlled [12]–[14].

Formally, the generality of the FTs can be attributed to the way probability distribution functions of particular observables behave under symmetry-breaking forcings, such as non-conservative and/or time-dependent forces (see [15, 16] for reviews on FT). Although the FTs are not expected to hold for every experimental system in contact with a thermal bath, they still provide a nice framework to analyse and understand the fluctuating heat interchange of out-of-equilibrium systems in general. Here we focus, as in most of the recent works in this field, on the family of systems whose driven dynamics can be well described by Langevin equations (without memory) or by discrete Markov chains. These are the prototypical models for which most FTs can be proven easily, without making any additional assumptions. (Note that fluctuation theorems have also been derived for non-Markovian dynamics [17, 18].)

FTs are exact relations for the probability distributions for the values W of observables $\mathcal{W}[\mathbf{x};\sigma]$ which are functionals of the stochastic state-space system trajectories $\mathbf{x} \equiv \{x(t)\}_{t=0}^{\tau}$ (e.g. work, heat or, more generally, different forms of trajectory-dependent entropy production) in processes driven by an arbitrary time-dependent protocol $\sigma(t)$ during a time τ. Typically, the so-called integral fluctuation theorems (IFTs) are exact relations for the thermal average over histories such as $\langle \mathrm{e}^{-\mathcal{W}} \rangle = 1$ while detailed fluctuation theorems (DFTs) for the same observable are stronger relations, typically of the form $P(W)/P^{\mathrm{T}}(-W) = \mathrm{e}^{W}$, involving the probability distributions $P(W) \equiv \langle \delta(\mathcal{W}[\mathbf{x};\sigma] - W) \rangle$ and $P^{\mathrm{T}}(W) \equiv \langle \delta(\mathcal{W}[\mathbf{x};\sigma] - W) \rangle^{\mathrm{T}}$ of the stochastic values W of $\mathcal{W}$, with T denoting that the trajectories x are sampled from a transformed (typically time-reversed) dynamics such that $\mathcal{W}^{\mathrm{T}} = -\mathcal{W}$.

The observable $\mathcal{W}$ is thus convenient to measure the asymmetry under the transformation T. Indeed, $\langle W \rangle \equiv D_{\mathrm{KL}}(P(W) \parallel P^{\mathrm{T}}(-W))$, with $D_{\mathrm{KL}}(P_A \parallel P_B)$ the Kullback–Leibler distance between two distributions P_A and P_B. Prominent examples for physically relevant observables $\mathcal{W}$ are the total entropy production $S_{\mathtt{tot}}$ which yields the Jarzynski IFT [7] and Crooks DFT [8], the non-adiabatic entropy S_{na} produced in transitions between non-equilibrium steady states yielding the Hatano–Sasa IFT [9] and DFT [20], and the adiabatic entropy production S_{a}, yielding the Speck–Seifert IFT [10] and DFT [19].

A simple unifying picture for all these seemingly different FTs has emerged recently [11, 19, 20]. Although a DFT trivially includes as a particular case an IFT it is now clear that there exist two basic operations T that we can use for generating the DFT version corresponding to each of the above-mentioned IFTs. These two operations are *time-reversal* (R), which reverses the protocol maintaining the form of the dynamical equations, and the transformation to the so-called *dual dynamics* (†), which corresponds to different dynamical equations. These transformations are interesting since they are closely related to symmetry operations of equilibrium and non-equilibrium steady states. The above-mentioned FTs can indeed be all unified in three detailed fluctuation theorems [19, 21, 22] satisfied by the total entropy production $S_{\mathtt{tot}}$, and by each term of its particular twofold decomposition, $S_{\mathtt{tot}} = S_{\mathrm{a}} + S_{\mathrm{na}}$, into an adiabatic S_{a} and a non-adiabatic part S_{na}. This splitting is physically motivated and closely related, for isothermal processes, to the splitting $Q_{\mathrm{tot}} = Q_{\mathrm{hk}} + Q_{\mathrm{exc}}$ of the total work done by all forces, into a house-keeping heat Q_{hk} and an excess heat Q_{exc}, proposed by Oono and Paniconi [23] and later formalized by Hatano–Sasa [9] for describing steady-state thermodynamics. The three DFT are

$$P(S_{\mathtt{tot}})/P^{\mathrm{R}}(-S_{\mathtt{tot}}) = \mathrm{e}^{S_{\mathtt{tot}}} \tag{1}$$

$$P(S_{\mathrm{na}})/P^{\dagger R}(-S_{\mathrm{na}}) = \mathrm{e}^{S_{\mathrm{na}}}, \tag{2}$$

$$P(S_{\mathrm{a}})/P^{\dagger}(-S_{\mathrm{a}}) = \mathrm{e}^{S_{\mathrm{a}}} \tag{3}$$

where in the second equation the dual and the time-reversal operations are composed. As we can see, at variance with the DFT for $S_{\mathtt{tot}}$, where the distribution of $S_{\mathtt{tot}}$ for the forward process is compared with the one for the backward process obeying a time-reversed protocol, the detailed fluctuation theorems for S_{a} and S_{na} require comparing the forward process with a process governed by the adjoint dynamics (which can be additionally time-reversed), which is very different from the original physical dynamics and therefore difficult to impose experimentally in general.

In a recent paper we have shown that writing detailed theorems in terms of joint probabilities is a convenient approach for deriving easily a family of relations, including the three detailed FTs, which might prove relevant for diverse applications. Introducing joint distributions is, on the other hand, natural when we are interested in the path-dependent fluctuations of non-scalar observables. Indeed more general transformations T such as vector rotations [24] can also be considered for deriving joint-distribution FTs. In this work we provide more detailed calculations and expand the results of [25], and then propose new applications.

The organization of this paper is as follows. In section 2 we review various FTs and introduce the notations and basic observables used in the next sections. In section 2.4 we review the main results of our previous work and give their detailed derivation using Langevin dynamics in its path-integral (Onsager–Machlup) representation (section 3.1) and Markov chains (section 3.2). In section 4 we discuss some applications of our results. Conclusions and perspectives are gathered in section 5.

2. Fluctuation theorem preliminaries

It is clear that, given two different dynamical weights $\mathcal{P}[\mathbf{x};\sigma]$ and $\mathcal{P}^{\mathrm{T}}[\mathbf{x};\sigma]$ for the stochastic trajectories $\mathbf{x} \equiv \{x(t)\}_{t=0}^{\tau}$ of a given system, we can always define a trajectory-dependent quantity

$$\mathcal{W}[\mathbf{x};\sigma] = \ln(\mathcal{P}[\mathbf{x};\sigma]/\mathcal{P}^{\mathrm{T}}[\mathbf{x};\sigma]). \tag{4}$$

which satisfies, by construction, the symmetry relation

$$\langle \mathcal{O}[\mathbf{x};\sigma]\rangle_{p_0} = \langle \mathcal{O}[\mathbf{x};\sigma]\mathrm{e}^{\mathcal{W}[\mathbf{x};\sigma]}\rangle_{p_1}^{\mathrm{T}}, \tag{5}$$

where $\langle\cdots\rangle = \int \mathcal{D}\mathbf{x}\mathcal{P}[\mathbf{x};\sigma]\ldots$, $\langle\cdots\rangle^{\mathrm{T}} = \int \mathcal{D}\mathbf{x}\mathcal{P}^{\mathrm{T}}[\mathbf{x};\sigma]\ldots$, with initial conditions sampled from the arbitrary distributions $p_0(x)$ and $p_1(x)$, respectively. The functional $\mathcal{O}[\mathbf{x},\sigma]$ is an arbitrary observable and thus equation (5) can be used to get a detailed statistics. $\mathcal{W}$

is, by construction, odd under the swapping of the two dynamics and its average over the first dynamics is the Kullback–Leibler distance between the two trajectory ensembles $\mathcal{P}[\mathbf{x};\sigma]$ and $\mathcal{P}^{\mathrm{T}}[\mathbf{x};\sigma]$, that is $\langle\mathcal{W}\rangle = \mathcal{D}_{\mathrm{KL}}(\mathcal{P}[\mathbf{x};\sigma] \parallel \mathcal{P}^{\mathrm{T}}[\mathbf{x};\sigma])$. One may use the general equation (5) in two interesting ways.

- Mapping trajectory ensembles. On the one hand we might directly choose two different dynamics (through their transition probabilities or through their Langevin equations, for instance) and use the equation to compare them quantitatively and from a purely information theoretical point of view. At this respect it is worth noting equation (5) 'maps' averages of arbitrary observables $\mathcal{O}$ in the original trajectory ensemble to another average in a 'target ensemble' of trajectories. An interesting case is the one which maps averages in a true non-equilibrium ensemble where detailed balance is broken (e.g. driven by non-conservative forces) to one satisfying detailed balance (e.g. driven by potential forces). A particular instance where $\mathcal{W}$ does also acquire a clear physical meaning is the mapping to equilibrium dynamics that is discussed in section 4.2.3.
- Symmetry transformations and entropy production. On the other hand, instead of choosing a 'target' dynamics, we might focus on the properties of the original dynamics and directly choose transformations T of it connected to symmetries of their equilibrium and non-equilibrium steady states (NESS). As is well known and we briefly review in the next sections, what makes this case specially interesting is that for such special transformations usually $\mathcal{W}$ acquires a well-defined physical meaning as generalized trajectory entropy productions. As we show in the following sections, time-reversal and the transformation to the so-called 'dual' dynamics yield the three detailed theorems of (3) and thus many known fluctuation theorems of great relevance in stochastic thermodynamics.

2.1. Time-reversed dynamics

In the *time-reversed* dynamics, $T = R$, trajectories are still governed by the original dynamical equations, with a time-reversed protocol $\sigma^{\mathrm{R}}(t) = \sigma(\tau - t)$. When a system is driven out of equilibrium, either by non-conservative or time-dependent forces, the resulting asymmetry is, for example, measured by comparing the statistical weights of two 'twin trajectories' [20] $\mathbf{x}$ and $\mathbf{x}^{\mathrm{R}}$ such that their components are related as $x^{\mathrm{R}}(t) = x(\tau - t)$, that is $\mathbf{x}$ is evolved with $\sigma(t)$ and $\mathbf{x}^{\mathrm{R}}$ with $\sigma^{\mathrm{R}}(t)$. A convenient measure between the two weights is the functional

$$\mathcal{S}[\mathbf{x};\sigma] = \ln(\mathcal{P}[\mathbf{x};\sigma]/\mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]), \qquad (6)$$

defined from the statistical weight $\mathcal{P}[\mathbf{x};\sigma]$ of system trajectories $\mathbf{x}$ evolved, starting from an initial condition distribution $\mathrm{Prob}[x(0) = s] = p_0(s)$, during a given time interval τ in a d-dimensional state space under the action of forces controlled by an arbitrary set $\sigma \equiv \sigma(t)$ of time-dependent external parameters [5, 6, 11]. Here $\mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]$ denotes the statistical weight of the trajectory $\mathbf{x}$ but evolved backwards, $x^{\mathrm{R}}(t) = x(t - \tau)$, and with the backward protocol $\sigma^{\mathrm{R}}(t) = \sigma(\tau - t)$, sampled from an initial condition distribution $\mathrm{Prob}[x^{\mathrm{R}}(0) = s] = p_1(s)$. For instance, for a Langevin dynamics we typically have $\mathcal{P}[\mathbf{x};\sigma] \sim \mathrm{e}^{-\int_0^\tau \mathrm{d}t\,\mathcal{L}[x,\dot{x};\sigma] + \ln p_0(x(0))}$ and $\mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}] \sim \mathrm{e}^{-\int_0^\tau \mathrm{d}t\,\mathcal{L}[x^{\mathrm{R}},\dot{x}^{\mathrm{R}};\sigma^{\mathrm{R}}] + \ln p_1(x^{\mathrm{R}}(0))}$, where $\mathcal{L}[x,\dot{x};\sigma]$ is the dynamical action of the system which, in general, also contains the

J. Stat. Mech. (2012) P02009

logarithm of a Jacobian, whose form depends on the considered stochastic calculus. By making a simple change of variables, we easily derive $\mathcal{S}[\mathbf{x};\sigma] \sim -\int_0^\tau \mathrm{d}t\,(\mathcal{L}[x,\dot{x};\sigma] - \mathcal{L}[x,-\dot{x};\sigma]) - \ln(p_0(x(0))/p_1(x(\tau)))$ (see section 3.1). From (6) we have, by construction, that $\mathcal{S}$ is odd under $T = R$, that is, $\mathcal{S}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}] = -\mathcal{S}[\mathbf{x};\sigma]$, since R is an involution. We also note that the definition of $\mathcal{S}$ allows us to write, for the average of any observable $\mathcal{O}[\mathbf{x};\sigma]$

$$\langle \mathcal{O}[\mathbf{x};\sigma]\rangle = \langle \mathcal{O}[\mathbf{x}^{\mathrm{R}};\sigma]\mathrm{e}^{-\mathcal{S}[\mathbf{x};\sigma^{\mathrm{R}}]}\rangle^{\mathrm{R}}, \tag{7}$$

where $\langle\cdots\rangle = \int \mathcal{D}\mathbf{x}\mathcal{P}[\mathbf{x};\sigma]\cdots$ and $\langle\cdots\rangle^{R} = \int \mathcal{D}\mathbf{x}\mathcal{P}^{\mathrm{R}}[\mathbf{x};\sigma^{\mathrm{R}}]\cdots$ denote the average over forward and reversed trajectories, respectively. We note also that the concept of twin trajectories used to define $\mathcal{S}$ is irrelevant in the last equation since on each side we integrate over all possible trajectories (i.e. $\mathbf{x}$ is now a dummy variable) and that $\int \mathcal{D}\mathbf{x}^{\mathrm{R}} = \int \mathcal{D}\mathbf{x}$ since x^{R} is simply a time reflection and shift of the trajectory $\mathbf{x}$ in time-state space. Equation (7) implies, in particular,

$$\langle \mathrm{e}^{-\lambda\mathcal{S}[\mathbf{x};\sigma]}\rangle = \langle \mathrm{e}^{-(1-\lambda)\mathcal{S}[\mathbf{x};\sigma^{\mathrm{R}}]}\rangle^{\mathrm{R}}, \tag{8}$$

where on the left-hand side we recognize the generating function of $P(S) = \langle\delta(\mathcal{S}[\mathbf{x};\sigma]-S)\rangle$ (we use calligraphic symbols to differentiate functionals of stochastic trajectories from their actual values) with λ an arbitrary number we can use to compute any cumulant of S. Introducing $P^{\mathrm{R}}(S) = \langle\delta(\mathcal{S}[\mathbf{x};\sigma^{\mathrm{R}}] - S)\rangle^{\mathrm{R}}$, it is then straightforward to derive the detailed FT (DFT):

$$P(S)/P^{\mathrm{R}}(-S) = \mathrm{e}^{S} \tag{9}$$

which implies, by direct integration or by setting $\lambda = 1$ above, the integral FT (IFT)

$$\langle \mathrm{e}^{-\mathcal{S}[\mathbf{x};\sigma]}\rangle = 1 \tag{10}$$

and thus, by using Jensen's inequality, we get

$$\langle \mathcal{S}[\mathbf{x};\sigma]\rangle \geq 0. \tag{11}$$

This inequality can also be obtained by noting that $\langle\mathcal{S}\rangle$ is equal to the positively defined Kullback–Leibler distance between two probability distributions, that is, for the present case, $\langle\mathcal{S}\rangle = \int \mathcal{D}\mathbf{x}\ \mathcal{P}[\mathbf{x};\sigma]\mathcal{S}[\mathbf{x};\sigma] = \mathcal{D}_{\mathrm{KL}}(\mathcal{P}[\mathbf{x};\sigma] \parallel \mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]) \geq 0$. Since the equality above is thus reached for time-reversal symmetric processes, at equilibrium $\langle\mathcal{S}\rangle = 0$, as expected. It is easy to show that such symmetry is actually stronger, $\mathcal{S} = 0$, due to detailed balance. In addition, if the relaxation time of the equilibrium states is finite, then processes that are driven very slowly compared with it do not produce this quantity either, as the adiabatic process makes the system visit a sequence of the (symmetric) equilibrium states compatible with the instantaneous value $\sigma(t)$. It is also worth remarking here that all the above statistical properties are valid for arbitrary protocols $\sigma(t)$, arbitrary initial conditions for the forward p_0 and backward p_1 process, and they are time-independent, that is, they are valid for any $\tau \geq 0$. Let us note, however, that our definition of $\mathcal{S}$ includes a border term containing information about the initial condition, p_0 and p_1, of the twin processes it compares. We note also that the choice of p_1 is free, and not a constraint for the trajectories at the border $t = \tau$ however, a direct connection with 'physical' entropy production $S_{\mathtt{tot}}$ can be done if one chooses p_1 being the solution of the Fokker–Planck (or master) equation at time τ [11].

J. Stat. Mech. (2012) P02009

What makes $\mathcal{S}$ physically interesting in connection with stochastic thermodynamics is that it can be identified, up to a time border term, with the stochastic heat dissipated into the reservoir divided by its temperature along the stochastic system trajectory x, for an isothermal process. If such process is described, for instance, by a Langevin dynamics $\dot{x} = f(x;\sigma) + \xi$, where the force $f(x;\sigma)$ might contain both conservative and non-conservative terms, it can be shown that [5, 26] (see sections 3.1 and 3.2 for the derivation in Langevin dynamics and Markov chain situations, respectively)

$$\mathcal{S}[\mathbf{x};\sigma] = \frac{1}{T}\int_0^\tau \mathrm{d}t\ f(x;\sigma)\dot{x} + \ln\frac{p_0(x(0))}{p_1(x(\tau))} \tag{12}$$

$$= \frac{Q^{\mathtt{tot}}[\mathbf{x};\sigma]}{T} + \ln\frac{p_0(x(0))}{p_1(x(\tau))} \tag{13}$$

where the functional of the trajectory $Q^{\mathtt{tot}}[\mathbf{x};\sigma]$ is the total work done by the force $f(x;\sigma)$ on the stochastic trajectory x under the protocol σ, and the border term $\ln[p_0(x(0))/p_1(x(\tau))]$ depends only on the time boundaries of each stochastic trajectory, sampled by p_0 and p_1 which remain so far arbitrary. With this identification of $\mathcal{S}$, it is also interesting to note again that, since $\langle\mathcal{S}\rangle = \mathcal{D}_{\mathrm{KL}}(\mathcal{P}[\mathbf{x};\sigma] \parallel \mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}])$, (11) and (12) relate irreversibility and dissipation in an elegant information theoretical way. In this respect it is worth noting that in equilibrium σ is constant, $p_{0,1}(x) = \mathrm{e}^{-\beta(E(x;\sigma)-F(\sigma))}$, and since all forces are conservative, we have $Q^{\mathtt{tot}} = E[x(0),\sigma] - E[x(\tau),\sigma]$, yielding $\mathcal{S}[\mathbf{x};\sigma] = 0$ for each stochastic trajectory $\mathbf{x}$. By writing $\mathcal{S}$ in terms of transition probabilities one obtains that this is equivalent to the detailed balance condition.

Equations (7–12) can be combined to derive many known FTs for Markovian systems. To start, by combining (7) and (12), one obtains the generalized Crooks relation for the average of an arbitrary observable $\mathcal{O}[\mathbf{x};\sigma]$ along a trajectory with the forward and backward protocol:

$$\langle\mathcal{O}[\mathbf{x};\sigma]\mathrm{e}^{-(1/T)Q^{\mathtt{tot}}[\mathbf{x};\sigma]-\ln(p_0(x(0);\sigma(0))/p_1(x(\tau);\sigma(\tau)))}\rangle_{p_0} = \langle\mathcal{O}[\mathbf{x};\sigma]\rangle^{\mathrm{R}}_{p_1}. \tag{14}$$

We now have a lot of freedom to derive FTs, by choosing appropriate values for $\mathcal{O}$, p_0 and p_1.

- *Seifert relation.* Choosing $\mathcal{O} = 1$ and $p_1(x) = \rho(x,\tau)$ to be the time-dependent solution of the Fokker–Planck equation with initial condition $p_0(x) \equiv \rho(x,0)$, we obtain the Seifert theorem [11], valid for all times τ and arbitrary initial conditions. This choice defines the so-called *trajectory-dependent total entropy production* $S_{\mathtt{tot}} = Q^{\mathtt{tot}}/T + S_{\mathrm{s}}$, with $S_{\mathrm{s}} = \ln[\rho(x,\tau)/\rho(x,0)]$ the *trajectory-dependent system entropy production*, such that

$$\left\langle \mathrm{e}^{-S_{\mathtt{tot}}}\right\rangle_{\text{any } p_0} = 1. \tag{15}$$

Seifert also made the interesting observation that this choice for p_1 is optimal in the sense $\min_{p_1}[\langle\mathcal{S}\rangle_{p_0}] = \langle\mathcal{S}_{\mathrm{tot}}\rangle_{p_0}$. This is easy to understand if we write

$$\langle\mathcal{S} - \mathcal{S}_{\mathrm{tot}}\rangle_{p_0} = \left\langle \ln\frac{\rho(x(\tau),\tau)}{p_1(x(\tau))}\right\rangle_{p_0}$$
$$= \int \mathrm{d}y\, \rho(y,\tau)\ln\frac{\rho(y,\tau)}{p_1(y)} = \mathcal{D}_{\mathrm{KL}}(\rho(y,\tau) \parallel p_1(y)) \geq 0, \tag{16}$$

J. Stat. Mech. (2012) P02009

and use the positiveness of the Kullback–Leibler distance between any two distributions.

- *Jarzynski relation for NESS.* By choosing $\mathcal{O}=1$, $p_1(x)=\rho_{\rm SS}(x;\sigma(\tau))$ and $p_0(x)=\rho_{\rm SS}(x;\sigma(0))$ from the steady-state solution of the Fokker–Planck equation $\rho_{\rm SS}(x;\sigma)$ we obtain the generalized Jarzynski relation for transitions between NESS, valid for all times τ and steady-state initial conditions compatible with $\sigma(0)$:

$$\left\langle \mathrm{e}^{-(Q^{\mathtt{tot}}[\mathbf{x};\sigma]/T)-[\phi(x(\tau),\sigma(\tau))-\phi(x(0);\sigma(0))]} \right\rangle_{\rho_{\rm SS}(\sigma(0))} = 1 \qquad (17)$$

where $\phi(x;\sigma)\equiv -\ln\rho_{\rm SS}(x;\sigma)$. As noted by Hatano and Sasa this relation does not generalize the second law of thermodynamics for NESS, since its corresponding Jensen's inequality:

$$\left\langle \frac{Q^{\mathtt{tot}}[\mathbf{x};\sigma]}{T} + [\phi(x(\tau),\sigma(\tau))-\phi(x(0);\sigma(0))] \right\rangle_{\rho_{\rm SS}(\sigma(0))} \geq 0, \qquad (18)$$

does not reach zero for adiabatic processes due to the presence of non-conservative forces which, even in the steady state, inject energy and produce the so-called house-keeping heat.

- *Jarzynski relation.* The IFT of equation (17) reduces to the well-known Jarzynski relation if the initial steady-state is the Boltzmann–Gibbs equilibrium state $\rho_{\rm SS}(x;\sigma)=\rho_{\rm eq}(x;\sigma)=\mathrm{e}^{-\beta[E(x;\sigma)-F(\sigma)]}$, with $E(x;\sigma)$ the energy of state x and $F(\sigma)$ the free energy, under the constraint σ:

$$\left\langle \mathrm{e}^{-\beta\mathcal{W}_{\rm d}} \right\rangle_{\rho_{\rm eq}(\sigma(0))} = \mathrm{e}^{-\beta\Delta F} \qquad (19)$$

where the dissipated work is $\mathcal{W}_{\rm d}[\mathbf{x};\sigma]\equiv Q^{\mathtt{tot}}[x;\sigma]+E[x(\tau);x(\tau)]-E[x(0);\sigma(0)]$ and $\Delta F=F(\sigma(\tau))-F(\sigma(0))$. Again

$$\langle \mathcal{W}_{\rm d}\rangle_{\rho_{\rm eq}(\sigma(0))} \geq \Delta F. \qquad (20)$$

The equality is achieved in the adiabatic limit and thus Jensen's inequality yields the second law for transitions between equilibrium states. Note that (an infinite number) of other second-law-like inequalities can be obtained from variational methods [28].

- *Crooks relation.* If $p_0(x)=\rho_{\rm eq}(x;\sigma(0))$ and $p_1(x)=\rho_{\rm eq}(x;\sigma(\tau))$, and $\mathcal{O}[\mathbf{x};\sigma]=\delta(\mathcal{W}_{\rm d}[\mathbf{x};\sigma]-W_{\rm d})$, equation (7) reduces to

$$P(W_{\rm d})\mathrm{e}^{-\beta\Delta F} = P^{\rm R}(-W_{\rm d})\mathrm{e}^{-W_{\rm d}}$$

where $P(W_{\rm d})=\langle\delta(\mathcal{W}_{\rm d}[\mathbf{x};\sigma]-W_{\rm d})\rangle$ and $P^{\rm R}(W_{\rm d})=\langle\delta(\mathcal{W}^{\rm R}_{\rm d}[\mathbf{x};\sigma]-W_{\rm d})\rangle^{\rm R}$,

- *Fluctuation in NESS.* In the absence of time-dependent forces, $\sigma(t)=\sigma_0$, a system initially prepared in the steady state $p_0(x)=\rho_{\rm SS}(x;\sigma_0)$ remains there and satisfies a particular form of the Crooks relation:

$$P(W_{\rm d}) = P(-W_{\rm d})\mathrm{e}^{-W_{\rm d}}. \qquad (21)$$

J. Stat. Mech. (2012) P02009

2.2. Dual dynamics

Non-equilibrium steady states (NESS) are already asymmetric with respect to time-reversal due to the lack of detailed balance. This has led Oono and Paniconi [23] to introduce a useful twofold decomposition of the total heat exchange into a 'house-keeping heat' part, constantly produced to maintain the non-equilibrium driven steady state with non-vanishing currents, and an 'excess heat' part, produced only during transitions between steady states. The excess heat is minimized in adiabatic processes, that is when $\sigma(t)$ varies very slowly compared with an assumed finite relaxation time towards the NESS, while the house-keeping heat is minimized only at equilibrium, in the absence of non-conservative forces, when detailed balance is recovered. Hatano and Sasa formalized this splitting for Langevin dynamics by defining stochastic trajectory-dependent quantities, $Q^{\mathrm{hk}}[\mathbf{x};\sigma]$ and $Q^{\mathrm{ex}}[\mathbf{x};\sigma]$, for the house-keeping and excess heat, respectively, and deriving an IFT which generalized the second law for transitions between NESS. Although they do not use it for its derivation, they point out that the so-called dual dynamics, denoted by the symbol †, composed with time-reversal, plays a role analogous to time-reversal alone in the derivation of the Jarzynski equality.

Indeed, the adjoint transformation is defined such that NESS are symmetric with respect to the simultaneous application of time-reversal R and the adjoint transformation † to the original dynamics, that is, with respect to the composed '$T = R\circ\dagger$' transformation: the steady state of the adjoint dynamics has the same distribution $\rho^{\dagger}_{\mathrm{SS}} = \rho_{\mathrm{SS}}$ as the original dynamics but with an inverted steady-state current $J^{\dagger}_{\mathrm{SS}} = -J_{\mathrm{SS}}$. Since the original current can be recovered by a time-reversal without changing ρ_{SS}, the steady state is symmetric with respect to the composed operation $T = R\circ\dagger$. It is worth remarking that the adjoint transformation changes the dynamics (see section 3.1.2 for an explicit example). It is then natural to introduce the so-called trajectory-dependent non-adiabatic entropy production

$$\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma] = \ln(\mathcal{P}[\mathbf{x};\sigma]/\mathcal{P}^{\dagger R}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]), \tag{22}$$

to measure the asymmetry produced when the system is driven out of the NESS. Here $\mathcal{P}^{\dagger R}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]$ is the weight of the trajectory $\mathbf{x}$ in the time-reversed dual dynamics. In section 2.4 we derive explicit forms for S_{na} for Langevin and Markov chain dynamics. For Langevin dynamics we get

$$\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma] = -\int_0^{\tau} \mathrm{d}t\, \frac{\partial\phi}{\partial x}(x;\sigma)\dot{x} + \ln\frac{p_0(x(0))}{p_1(x(\tau))} \tag{23}$$

$$= \frac{Q^{\mathrm{ex}}[\mathbf{x};\sigma]}{T} + \ln\frac{p_0(x(0))}{p_1(x(\tau))} \tag{24}$$

making a physical connection with the Oono–Paniconi–Hatano–Sasa excess heat. Note that here we have used again the freedom to choose the initial condition distributions p_0 and p_1 for the 'twin trajectories', the first weighted in the forward protocol of the physical dynamics and the second weighted in the backward protocol of the dual dynamics. By analogy, it is now straightforward to write FTs. The definition of $\mathcal{S}_{\mathrm{na}}$ allows us to write, for the average of any observable $\mathcal{O}[\mathbf{x};\sigma]$, the FT generator

$$\langle \mathcal{O}[\mathbf{x};\sigma]\mathrm{e}^{-\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]}\rangle_{p_0} = \langle \mathcal{O}[\mathbf{x};\sigma]\rangle^{R\dagger}_{p_1}, \tag{25}$$

J. Stat. Mech. (2012) P02009

where $\langle\cdots\rangle^{\dagger R} = \int \mathcal{D}\mathbf{x}\mathcal{P}^{\dagger R}[\mathbf{x};\sigma^{\mathrm{R}}]\cdots$ denotes the average over trajectories generated by the dual dynamics and controlled by a time-reversed protocol and $\mathcal{S}^{\dagger}_{\mathrm{na}}[\mathbf{x}^{R};\sigma^{\mathrm{R}}] = -\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]$. It is worth remarking here that the above statistical relation is valid for any initial condition distributions p_0 and p_1 and is valid for any $\tau \geq 0$. We can now easily derive several relations by making particular choices for $\mathcal{O}$ and $p_{0,1}$.

- *Generating function*: By choosing $\mathcal{O}[\mathbf{x};\sigma] = \mathrm{e}^{(1-\lambda)\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]}$, with λ any number, we get
$$\langle \mathrm{e}^{-\lambda\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]}\rangle_{p_0} = \langle \mathrm{e}^{-(1-\lambda)\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma^{\mathrm{R}}]}\rangle^{R\dagger}_{p_1}. \tag{26}$$
- *Non-adiabatic entropy production DFT*. By choosing $\mathcal{O}[\mathbf{x};\sigma] = \delta(\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma] - S)$ and defining $P(S) = \langle\delta(\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma] - S)\rangle_{p_0}$ and $P^{R\dagger}(S) = \langle\delta(\mathcal{S}^{\dagger R}_{\mathrm{na}}[\mathbf{x};\sigma] - S)\rangle^{R\dagger}_{p_1} = \langle\delta(\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma] + S)\rangle^{R\dagger}_{p_1}$, it is straightforward to derive the DFT [19, 20]:
$$P(S_{\mathrm{na}})/P^{\mathrm{R}\dagger}(-S_{\mathrm{na}}) = \mathrm{e}^{S_{\mathrm{na}}} \tag{27}$$
which is the detailed version of the Hatano–Sasa IFT if the initial condition is the stationary distribution.
- *Hatano–Sasa IFT*. By choosing $\mathcal{O}[\mathbf{x};\sigma] = 1$ and $p_0(x) = \rho_{\mathsf{ss}}(x;\sigma(0))$ and $p_1(x) = \rho_{\mathsf{ss}}(x;\sigma(\tau))$ we obtain the Hatano–Sasa IFT:
$$\langle \mathrm{e}^{-(\beta Q^{\mathbf{ex}}[\mathbf{x};\sigma]+\Delta\phi)}\rangle_{\rho_{\mathsf{ss}}(\sigma(0))} = 1 \tag{28}$$
where $\Delta\phi = \phi(x(\tau);\sigma(\tau)) - \phi(x(0);\sigma(0))$. By using Jensen's inequality we get
$$\beta\langle Q^{\mathbf{ex}}[\mathbf{x};\sigma]\rangle_{\rho_{\mathsf{ss}}(\sigma(0))} \geq -\langle\Delta\phi\rangle_{\rho_{\mathsf{ss}}(\sigma(0))} \tag{29}$$
which is the generalization of the second law for the transition between NESS. Again, this inequality can also be obtained by noting that $\langle\mathcal{S}_{\mathrm{na}}\rangle$ is equal to the positively defined Kullback–Leibler distance between two probability distributions. For the present case, $\langle\mathcal{S}\rangle = \mathcal{D}_{\mathrm{KL}}(\mathcal{P}[\mathbf{x};\sigma] \parallel \mathcal{P}^{\dagger R}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]) \geq 0$, and the equality is reached for processes that are time-reversal symmetric in the dual dynamics. Therefore equilibrium in particular, and NESS in general, have $\langle\mathcal{S}\rangle = 0$, as expected. The latter is also true for adiabatic processes, slow compared to an assumed finite relaxation time towards the NESS, so that the system is always very close to the NESS corresponding to the instantaneous value of $\sigma(t)$. The absence of non-adiabatic entropy production in NESS is actually 'detailed': $\mathcal{S}_{\mathrm{na}} = 0$. As shown in section 3.1.2 this can be understood from the definition of dual dynamics and from a detailed balance-like relation between the transition probabilities of the direct and dual dynamics for the same pair of states.

Finally, we note that equilibrium states are symmetric with respect to R and $\dagger$. It is thus natural to introduce a new quantity $\mathcal{S}_{\mathrm{a}}$ to measure the asymmetry produced by non-conservative or time-dependent driving forces by using the $\dagger$ operation alone:
$$\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma] = \ln(\mathcal{P}[\mathbf{x};\sigma]/\mathcal{P}^{\dagger}[\mathbf{x};\sigma]) \tag{30}$$
where we note that the twin trajectories are actually the same, the first weighted in the direct dynamics and the second in the dual dynamics. In section 2.4 we derive an explicit form for S_{a}, from Langevin and Markov chain dynamics. For Langevin dynamics the

following relation holds:

$$\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma] = \frac{1}{T}\int_0^\tau \mathrm{d}t\, v_{\mathrm{SS}}(x;\sigma)\dot{x} + \ln\frac{p_0(x(0))}{p_1(x(0))} \tag{31}$$

$$= \frac{Q^{\mathrm{hk}}[\mathbf{x};\sigma]}{T} + \ln\frac{p_0(x(0))}{p_1(x(0))} \tag{32}$$

where the steady-state velocity $v_{\mathrm{SS}}(x;\sigma) \equiv J_{\mathrm{SS}}(\sigma)/\rho_{\mathrm{SS}}(x;\sigma)$, with $\rho_{\mathrm{SS}}(x;\sigma) \equiv \mathrm{e}^{-\phi(x;\sigma)}$ the NESS probability distribution and J_{SS} the probability current in state space. We also note that the initial condition for trajectories weighted in the dual dynamics p_1 is not reversed. The above makes the physical connection with the house-keeping heat. By analogy with the previous cases, it is now straightforward to write a FT generator for this observable:

$$\langle \mathcal{O}[\mathbf{x};\sigma]\mathrm{e}^{-\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma]}\rangle_{p_0} = \langle \mathcal{O}[\mathbf{x};\sigma]\rangle^{\dagger}_{p_1}, \tag{33}$$

where $\langle\cdots\rangle^{\dagger} = \int \mathcal{D}x\,\mathcal{P}^{\dagger}[\mathbf{x};\sigma]\cdots$ denotes the average over trajectories weighted in the dual dynamics. We can thus proceed analogously.

- *Generating function.*

$$\langle \mathrm{e}^{-\lambda\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma]}\rangle_{p_0} = \langle \mathrm{e}^{-(1-\lambda)\mathcal{S}^{\dagger}_{\mathrm{a}}[\mathbf{x};\sigma]}\rangle^{\dagger}_{p_1}. \tag{34}$$

- *Adiabatic entropy production DFT.*

$$P(S_{\mathrm{a}})/P^{\dagger}(-S_{\mathrm{a}}) = \mathrm{e}^{S_{\mathrm{a}}} \tag{35}$$

  which is a detailed version of the Speck–Seifert IFT if the initial condition is stationary.

- *Speck–Seifert IFT.* By choosing $\mathcal{O}[\mathbf{x};\sigma] = 1$ and $p_0(x) = p_1(x) = \rho_{\mathrm{ss}}(x;\sigma(0))$ we obtain the Speck–Seifert IFT:

$$\langle \mathrm{e}^{-\beta Q^{\mathrm{hk}}[\mathbf{x};\sigma]}\rangle_{\rho_{\mathrm{ss}}(x;\sigma(0))} = 1 \tag{36}$$

  by using Jensen's inequality we get

$$\langle \beta Q^{\mathrm{hk}}[\mathbf{x};\sigma]\rangle_{\rho_{\mathrm{ss}}(x;\sigma(0))} \geq 0. \tag{37}$$

  Again, this inequality can also be obtained by noting that $\langle\mathcal{S}_{\mathrm{a}}\rangle$ is equal to the positively defined Kullback–Leibler distance between two probability distributions, that is, for the present case $\langle\mathcal{S}_{\mathrm{a}}\rangle = \mathcal{D}_{\mathrm{KL}}(\mathcal{P}[\mathbf{x};\sigma] \parallel \mathcal{P}^{\dagger}[\mathbf{x};\sigma]) \geq 0$. Since the equality above is reached for dual-symmetric processes by construction, only equilibrium states have $\langle Q^{\mathrm{hk}}\rangle = 0$. NESS do produce Q^{hk} because they need house-keeping energy to maintain detailed balance violation. In other words, NESS are dual asymmetric because the $\dagger$ operation, although keeping ρ_{ss} invariant, invert the steady-state current J_{SS}. Therefore, the equality in equation (37) is never reached by NESS. At equilibrium, the equality is actually reached and is 'detailed' in the sense that $Q^{\mathrm{hk}}[\mathbf{x};\sigma] = 0$ for each trajectory, since $J_{\mathrm{SS}} = 0$ exactly.

J. Stat. Mech. (2012) P02009

2.3. Splitting

By using the operations R and $\dagger$, related to symmetries of equilibrium and NESS states, we have defined the trajectory-dependent total, non-adiabatic and adiabatic entropy production functionals:

$$\mathcal{P}[\mathbf{x};\sigma] = \mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]\mathrm{e}^{\mathcal{S}[\mathbf{x};\sigma]} \tag{38}$$

$$\mathcal{P}[\mathbf{x};\sigma] = \mathcal{P}^{\dagger R}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]\mathrm{e}^{\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]} \tag{39}$$

$$\mathcal{P}[\mathbf{x};\sigma] = \mathcal{P}^{\dagger}[\mathbf{x};\sigma]\mathrm{e}^{\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma]}. \tag{40}$$

From the last equation we have

$$\mathcal{P}^{\dagger R}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}] = \mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]\mathrm{e}^{\mathcal{S}_{\mathrm{a}}^{\dagger}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]} \tag{41}$$

so therefore from the second we get

$$\mathcal{P}[\mathbf{x};\sigma] = \mathcal{P}^{\mathrm{R}}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]\mathrm{e}^{\mathcal{S}_{\mathrm{a}}^{\dagger}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}]+\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]}. \tag{42}$$

Comparing with the first we conclude:

$$\mathcal{S}[\mathbf{x};\sigma] = \mathcal{S}_{\mathrm{a}}^{\dagger}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}] + \mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma] \tag{43}$$

only using transformation properties. If we define $\mathcal{S}_{\mathrm{a}}$ such that it does not include a border term as is customary, we see that $\mathcal{S}_{\mathrm{a}}^{\dagger}[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}] = \mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma]$. This is so because R and $\dagger$ only change the sign of $\mathcal{S}_{\mathrm{a}}$. Therefore, we get

$$\mathcal{S} = \mathcal{S}_{\mathrm{a}} + \mathcal{S}_{\mathrm{na}} \tag{44}$$

which is the starting point of the three detailed theorems [19]. However, it is worth noting that in [19] $\mathcal{S}_{\mathtt{tot}}$ have been considered instead of $\mathcal{S}$. This splitting is explicitly obtained for Langevin and Markov chains in section 2.4.

2.4. Fluctuation theorems from joint distribution symmetries

In a previous work [25] we have shown that the use of joint probability distributions for different total entropy production decompositions is a convenient tool for deriving a variety of exact expressions for Markovian systems, including many known fluctuation theorems arising from particular twofold decompositions of the total entropy production. This was done by first noting that *any* decomposition of the total trajectory entropy production for Markov systems, $\mathcal{S}[\mathbf{x};\sigma] = \sum_{i=1}^{M}\mathcal{A}_i[\mathbf{x};\sigma]$, has a joint probability distribution satisfying a generalized detailed fluctuation theorem, when all the contributing terms are odd with respect to time-reversal, $\mathcal{A}_i[\mathbf{x}^{\mathrm{R}};\sigma^{\mathrm{R}}] = -\mathcal{A}_i[\mathbf{x};\sigma]$:

$$\frac{P(A_1, A_2, \ldots, A_M)}{P^{\mathrm{R}}(-A_1, -A_2, \ldots, -A_M)} = \mathrm{e}^{S} \tag{45}$$

with $S = \sum_{i=1}^{M} A_i$ and $P(A_1, A_2, \ldots, A_M) = \langle\delta(\mathcal{A}_1 - A_1)\cdots\delta(\mathcal{A}_M - A_M)\rangle_{p_0}$. This contains the same information as

$$\left\langle \mathrm{e}^{-\sum_{i=1}^{M}\lambda_i\mathcal{A}_i[\mathbf{x};\sigma]}\right\rangle_{p_0} = \left\langle \mathrm{e}^{-\sum_{i=1}^{M}(1-\lambda_i)\mathcal{A}_i[\mathbf{x};\sigma^{\mathrm{R}}]}\right\rangle_{p_1}^{\mathrm{R}}, \tag{46}$$

for $\{\lambda_i\}M$ arbitrary numbers. It is worth remarking that relations of the kind of the previous equation have been already derived in [6, 29] for time-independent σ.

More generally, for a transformation T, we can define a quantity $\mathcal{S}_{\mathrm{T}} = \ln\{\mathcal{P}[\mathbf{x};\sigma]/\mathcal{P}^{\mathrm{T}}[\mathbf{x};\sigma]\}$. If we can write $\mathcal{S}_{\mathrm{T}}[\mathbf{x};\sigma] = \sum_{i=1}^{M}\mathcal{B}_i[\mathbf{x};\sigma]$ such that each component is odd or even under T, $\mathcal{B}_i^{\mathrm{T}} = \epsilon_i^{\mathrm{T}}\mathcal{B}_i$ with $\epsilon_i^{\mathrm{T}} = \pm 1$, we get

$$\frac{P(B_1, B_2, \ldots, B_M)}{P^{\mathrm{T}}(\epsilon_1^{\mathrm{T}}B_1, \epsilon_2^{\mathrm{T}}B_2, \ldots, \epsilon_M^{\mathrm{T}}B_M)} = \mathrm{e}^{\mathcal{S}_{\mathrm{T}}} \tag{47}$$

or, equivalently,

$$\left\langle \mathrm{e}^{-\sum_{i=1}^{M}\lambda_i\mathcal{B}_i[\mathbf{x};\sigma]}\right\rangle_{p_0} = \left\langle \mathrm{e}^{\sum_{i=1}^{M}(1-\lambda_i)\epsilon_i^{\mathrm{T}}\mathcal{B}_i^{\mathrm{T}}[\mathbf{x};\sigma]}\right\rangle_{p_1}^{\mathrm{T}}. \tag{48}$$

On the other hand, for a not constrained list of variables $\{\mathcal{C}_i\}_{i=2}^{M}$, such that $\mathcal{C}_i^{\mathrm{T}} = \epsilon_i^{\mathrm{T}}\mathcal{C}_i$, we can also write

$$\left\langle \mathrm{e}^{-\lambda\mathcal{S}_{\mathrm{T}}+\sum_{i=2}^{M}\lambda_i\mathcal{C}_i[\mathbf{x};\sigma]}\right\rangle_{p_0} = \left\langle \mathrm{e}^{-(1-\lambda)\mathcal{S}_{\mathrm{T}}^{\mathrm{T}}+\sum_{i=2}^{M}\lambda_i\epsilon_i^{\mathrm{T}}\mathcal{C}_i^{\mathrm{T}}[\mathbf{x};\sigma]}\right\rangle_{p_1}^{\mathrm{T}}, \tag{49}$$

or, equivalently,

$$\frac{P(S_{\mathrm{T}}, C_2, C_3, \ldots, C_M)}{P^{\mathrm{T}}(-S_{\mathrm{T}}, \epsilon_2^{\mathrm{T}}C_2, \epsilon_3^{\mathrm{T}}C_3, \ldots, \epsilon_M^{\mathrm{T}}C_M)} = \mathrm{e}^{S_{\mathrm{T}}}. \tag{50}$$

These relations are interesting as they contain the three detailed fluctuation theorems [19] as a particular case for $M = 1$. For the three transformations $T = (R)$, $(\dagger)$, $(R \circ \dagger)$ we get, respectively, the DFTs for $\mathcal{S}$, $\mathcal{S}_{\mathrm{a}}$ and $\mathcal{S}_{\mathrm{na}}$. As an example, using the above relations for $M = 2$ we can get the following useful identities for $\mathcal{S}_{\mathrm{a}}$ and $\mathcal{S}_{\mathrm{na}}$:

$$P(\mathcal{S}_{\mathrm{a}}, \mathcal{S}_{\mathrm{na}}) = P^{\dagger}(-\mathcal{S}_{\mathrm{a}}, \mathcal{S}_{\mathrm{na}})\mathrm{e}^{\mathcal{S}_{\mathrm{a}}} = P^{\dagger R}(\mathcal{S}_{\mathrm{a}}, -\mathcal{S}_{\mathrm{na}})\mathrm{e}^{\mathcal{S}_{\mathrm{na}}} = P^{R}(-\mathcal{S}_{\mathrm{a}}, -\mathcal{S}_{\mathrm{na}})\mathrm{e}^{\mathcal{S}}. \tag{51}$$

The last one assumes the splitting $\mathcal{S} = \mathcal{S}_{\mathrm{a}} + \mathcal{S}_{\mathrm{na}}$.

Another interesting consequence of (50) and Bayes theorem is that

$$P(C_2, \ldots, C_M|S_{\mathrm{T}}) = P^{\mathrm{T}}(C_2^{\mathrm{T}}, \ldots, C_M^{\mathrm{T}}| - S_{\mathrm{T}}) \tag{52}$$

meaning that the variables $\{\mathcal{C}_i\}_{i=2}^{M}$ have identical statistical properties in the trajectory subensembles determined by the constraints $\mathcal{S}_{\mathrm{T}}[\mathbf{x},\sigma] = S_{\mathrm{T}}$ and $\mathcal{S}_{\mathrm{T}}^{\mathrm{T}}[\mathbf{x},\sigma] = -S_{\mathrm{T}}$, and can thus not be used to differentiate the original and the transformed dynamics. It is particularly instructive to consider variables $\{\mathcal{C}_i\}_{i=2}^{M}$ describing a discretized path of a certain duration. Following the discussion for the time-reversal case in [27] we can now write, for the particular transformations $T = (R)$, $(\dagger)$, $(R \circ \dagger)$, the subensembles equivalences

$$P(\mathrm{path}|S) = P^{\mathrm{R}}(\mathrm{path}^{\mathrm{R}}| - S) \tag{53}$$

$$P(\mathrm{path}|S_{\mathrm{a}}) = P^{\dagger}(\mathrm{path}| - S_{\mathrm{a}}) \tag{54}$$

$$P(\mathrm{path}|S_{\mathrm{na}}) = P^{\dagger R}(\mathrm{path}^{\mathrm{R}}| - S_{\mathrm{na}}). \tag{55}$$

As discussed in [27], and can be appreciated by recalling (3), the above expressions 'map' the statistics of typical trajectories in one forward process to rare trajectories in the transformed process, the latter being backward, the backward-dual or the forward-dual process, according to the corresponding transformation involved.

J. Stat. Mech. (2012) P02009

3. Markovian dynamics models

So far we have made mostly general mathematical considerations arising from the definitions of trajectory entropy production without specifying the dynamics behind the corresponding and different trajectory statistical weights. In this section we analyse two paradigmatic class of Markov dynamics. First we consider the Langevin dynamics for systems with configurations lying in the continuum and described by first-order stochastic differential equations. Second we consider the continuous-time Markov chains describing dynamical systems with discrete configurations, described by master equations with well-defined transition probabilities.

3.1. Langevin dynamics

3.1.1. Generalities. For simplicity we consider in the following discussion the case of a one-dimensional Langevin system coupled with a single thermal bath. Our model consists of a Brownian particle driven by an external force. In the presence of periodic boundary conditions, the steady state for this system has a non-zero probability current. Our results can be easily generalized to more dimensions and many particles (provided the temperature is the same along every spatial dimension—see, e.g., [58] for a physically relevant situation where it is not the case). We also consider in 3.2 the case of discrete Markov chains. We start with the Langevin equation

$$\dot{x} = -\frac{\partial U}{\partial x}(x;\alpha) + f + \xi, \tag{56}$$

where α represents a set of parameters of the system, f is a driving force and $\xi(t)$ is a Gaussian white noise with variance $\langle \xi(t)\xi(t')\rangle = 2T\delta(t-t')$. The parameters α and f may depend on time and $\sigma = (\alpha, f)$. For this system the probability of a given trajectory in the phase space has the following form:

$$\mathcal{P}[\mathbf{x};\sigma] = \int \mathcal{D}\xi P[\xi] p_0(x(0);\sigma_0) J[\mathbf{x};\sigma]\delta\left[\dot{x} + \frac{\partial U}{\partial x}(x;\alpha) - f - \xi\right], \tag{57}$$

where p_0 is the initial probability density function of the system (at $t = 0$), $P[\xi]$ is the probability distribution of the thermal noise and $J[\mathbf{x};\sigma]$ is a Jacobian to be defined below. $P[\xi]$ takes the following form:

$$P[\xi] = \Pi^{-1}(\tau)\exp\left[-\frac{1}{4T}\int_0^\tau \mathrm{d}t\,\xi(t)^2\right], \tag{58}$$

where $\Pi(\tau)$ is a normalization constant.

It is instructive to make at this point a formal, but important, remark. It is essential to specify the discretization scheme (i.e. the choice of Itō or Stratonovich stochastic calculus — see [30] for a detailed analysis) of the path integral given in equation (57) and all the following ones. In general, the relevant quantities we are interested in (different forms of heat) were originally defined in the Stratonovich scheme, and we work using this picture. Besides, the Stratonovich discretization is easier to use in our context since it is invariant under time-reversal, contrary to the Itō one. In the Stratonovich case, the

J. Stat. Mech. (2012) P02009

Jacobian in equation (57) is

$$J[\mathbf{x};\sigma] = \exp\left\{\frac{1}{2}\int_0^\tau \mathrm{d}t\,\frac{\partial^2 U}{\partial x^2}(x;\alpha)\right\}. \tag{59}$$

As regards time-reversal transformations, an important property of this Jacobian is the following:

$$J[\mathbf{x};\sigma] = J[\mathbf{x}_{\mathrm{R}};\sigma_R]. \tag{60}$$

Integrating out the noise in equation (57) and taking into account equation (58) we obtain

$$\mathcal{P}[\mathbf{x},\sigma] = \Pi^{-1}(\tau)J[\mathbf{x};\sigma]\exp\{-I[\mathbf{x};\sigma]\}, \tag{61}$$

where the Onsager–Machlup action functional $I[\mathbf{x};\sigma]$ is

$$I[\mathbf{x};\sigma] = -\ln p_0(x(0);\sigma_0) + \int_0^\tau \mathrm{d}t\left\{\frac{1}{4T}\left[\dot{x}^2 + \left(\frac{\partial U}{\partial x}(x;\alpha) - f\right)^2\right] + \frac{1}{2T}\dot{x}\left(\frac{\partial U}{\partial x}(x;\alpha) - f\right)\right\}. \tag{62}$$

Let us now consider the time-reversed probability weight given by

$$\mathcal{P}^{\mathrm{R}}[\mathbf{x}_{\mathrm{R}},\sigma_{\mathrm{R}}] = \Pi^{-1}(\tau)J[\mathbf{x}_{\mathrm{R}};\sigma_{\mathrm{R}}]\exp\{-I_{\mathrm{R}}[\mathbf{x}_{\mathrm{R}};\sigma_{\mathrm{R}}]\}, \tag{63}$$

with

$$I_{\mathrm{R}}[\mathbf{x}_{\mathrm{R}};\sigma_{\mathrm{R}}] = -\ln p_1(x(\tau);\sigma_\tau) + \int_0^\tau \mathrm{d}t\left\{\frac{1}{4T}\left[(\dot{x}_{\mathrm{R}})^2 + \left(\frac{\partial U}{\partial x}(x_{\mathrm{R}};\alpha_{\mathrm{R}}) - f_{\mathrm{R}}\right)^2\right] + \frac{1}{2T}\dot{x}_{\mathrm{R}}\left(\frac{\partial U}{\partial x}(x_{\mathrm{R}};\alpha_{\mathrm{R}}) - f_{\mathrm{R}}\right)\right\}, \tag{64}$$

where p_1 is, from now on, the solution of the Fokker–Planck equation of the process at time τ in order to make a connection with [11]. In this case one can immediately write for the total trajectory entropy production

$$\mathcal{S}[\mathbf{x};\sigma] = \ln(\mathcal{P}[\mathbf{x};\sigma]/\mathcal{P}^{\mathrm{R}}[\mathbf{x}_{\mathrm{R}};\sigma_{\mathrm{R}}]) = I_{\mathrm{R}}[\mathbf{x}_{\mathrm{R}};\sigma_{\mathrm{R}}] - I[\mathbf{x};\sigma] \equiv \mathcal{S}_{\mathrm{s}}[\mathbf{x};\sigma] + \mathcal{S}_{\mathrm{r}}[\mathbf{x};\sigma], \tag{65}$$

where one identifies the system (reservoir) entropy production $\mathcal{S}_{\mathrm{s}}[\mathbf{x};\sigma]$ ($\mathcal{S}_{\mathrm{r}}[\mathbf{x};\sigma]$) as

$$\mathcal{S}_{\mathrm{s}}[\mathbf{x};\sigma] = -\ln\frac{p_1(x(\tau);\sigma_\tau)}{p_0(x(0);\sigma_0)}, \tag{66}$$

$$\mathcal{S}_{\mathrm{r}}[\mathbf{x};\sigma] = -\frac{1}{T}\int_0^\tau \mathrm{d}t\,\dot{x}\left(\frac{\partial U}{\partial x}(x;\alpha) - f\right). \tag{67}$$

Consider now the instantaneous stationary probability density function for given values of the set of parameters of the system $\rho_{\mathrm{SS}}(x(t);\sigma(t)) = \mathrm{e}^{-\phi(x(t);\sigma(t))}$ and let us add and subtract the quantity $\int_0^\tau \mathrm{d}t\,\dot{x}(\partial\phi/\partial x)(x;\sigma)$ from equation (65). In this case one finds a different decomposition of the total trajectory entropy production in two different

J. Stat. Mech. (2012) P02009

contributions, the so-called adiabatic and the non-adiabatic contributions:

$$\mathcal{S}_{\mathrm{a}} = \int_0^\tau \mathrm{d}t\, \dot{x} \left[\frac{\partial\phi}{\partial x}(x;\sigma) - \frac{1}{T}\left(\frac{\partial U}{\partial x}(x;\alpha) - f\right)\right], \tag{68}$$

$$\mathcal{S}_{\mathrm{na}} = \ln \frac{p_0(x(0);\sigma_0)}{p_1(x(\tau);\sigma_\tau)} - \int_0^\tau \mathrm{d}t\, \dot{x} \frac{\partial\phi}{\partial x}(x;\sigma). \tag{69}$$

We thus have two relevant decompositions:

$$\mathcal{S}[\mathbf{x};\sigma] = \mathcal{S}_{\mathrm{s}}[\mathbf{x};\sigma] + \mathcal{S}_{\mathrm{r}}[\mathbf{x};\sigma] = \mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma] + \mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]. \tag{70}$$

Note that instead of using the Onsager–Machlup action functional $I[\mathbf{x};\sigma]$ one could equivalently work in the Martin–Siggia–Rose–Janssen–De Dominicis framework (which is more suitable for non-Markovian dissipation [18]).

3.1.2. Dual dynamics in continuous time. When the steady-state probability density function for a given system does not satisfy detailed balance, one usually introduces the dual dynamics in terms of its propagator. In a time-discretized picture one defines it via the relation

$$K^\dagger(x_i|x_{i+1};\sigma_i) = K(x_{i+1}|x_i;\sigma_i)\frac{\rho_{\mathrm{SS}}(x_i;\sigma_i)}{\rho_{\mathrm{SS}}(x_{i+1};\sigma_i)}. \tag{71}$$

As can be seen, when detailed balance holds the dual propagator is equal to the corresponding propagator for the dynamics of the system.

In the limit of continuous time we have that $x_i = x(s), \sigma_i = \sigma(s), x_{i+1} = x(s + \mathrm{d}s)$, $\mathrm{d}s \to 0$. In that case

$$K(x(s+\mathrm{d}s)|x(s);\sigma(s)) \simeq \exp\left[\frac{1}{2}\mathrm{d}s\frac{\partial^2 U}{\partial x^2}(x;\alpha) - \frac{1}{4T}\mathrm{d}s\left(\dot{x} + \frac{\partial U}{\partial x}(x;\alpha) - f\right)^2\right]. \tag{72}$$

From (71) and (72) we have

$$K^\dagger(x(s+\mathrm{d}s)|x(s);\sigma(s)) \simeq \exp\left[\frac{1}{2}\mathrm{d}s\frac{\partial^2 U}{\partial x^2}(x;\alpha) - \frac{1}{4T}\mathrm{d}s\left(\dot{x} - \frac{\partial U}{\partial x}(x;\alpha) + f\right)^2 \right.$$
$$\left. -\ \mathrm{d}s\,\dot{x}\frac{\partial\phi}{\partial x}(x;\sigma)\right]. \tag{73}$$

Now, from (73) we obtain for the general dual propagator:

$$K^\dagger(x,t|x',t';\sigma) = \Pi^{-1}(t-t')\int_{x(t')=x'}^{x(t)=x} Dx J[\mathbf{x};\sigma]\exp[-\mathcal{I}^\dagger[\mathbf{x};\sigma]], \tag{74}$$

where

$$\mathcal{I}^\dagger[\mathbf{x};\sigma] = \int_{t'}^{t}\mathrm{d}s\left\{\frac{\beta}{4}\left[\dot{x}(s) - \frac{\partial U}{\partial x}(x(s);\alpha(s)) + f(s)\right]^2 + \dot{x}\frac{\partial\phi}{\partial x}(x(s);\sigma(s))\right\} \tag{75}$$

and $J[\mathbf{x};\sigma]$ is given by equation (59). Before going further, let us write the action (75) in a clearer way. By adding and subtracting the quantity $(1/T)\dot{x}(\partial U/\partial x - f)$ in the integrand

J. Stat. Mech. (2012) P02009

of equation (75), we obtain

$$\mathcal{I}^\dagger[\mathbf{x};\sigma] = \int_{t'}^{t} \mathrm{d}s \left\{ \frac{1}{4T} \left[\dot{x}(s) + \frac{\partial U}{\partial x}(x(s);\alpha(s)) - f(s) \right]^2 \right\} + \mathcal{S}_\mathrm{a}[\mathbf{x};\sigma]. \quad (76)$$

We thus confirm the identification of $\mathcal{S}_\mathrm{a}$ with the trajectory relative entropy between its weight in the original dynamics with its weight in the dual dynamics, as defined in (30):

$$\mathcal{S}_\mathrm{a}[\mathbf{x};\sigma] = \mathcal{I}^\dagger[\mathbf{x};\sigma] - \mathcal{I}[\mathbf{x};\sigma] = \ln \frac{P[\mathbf{x};\sigma]}{P^\dagger[\mathbf{x};\sigma]}. \quad (77)$$

The second equality in the previous equation holds for $p_0(x(0)) = p_1(x(0))$. From this remark one obtains as an immediate result the validity of an IFT for the adiabatic entropy production:

$$\langle \mathrm{e}^{-\mathcal{S}_\mathrm{a}[\mathbf{x};\sigma]} \rangle = \langle 1 \rangle^\dagger \equiv 1. \quad (78)$$

To endow the dual dynamics with a physical meaning, one would like to associate the dual propagator with an effective Langevin equation (i.e. an effective microscopic dynamics). To do so, let us consider the following action:

$$\mathcal{I}_\mathrm{eff}[\mathbf{x},\sigma] = \int_{t'}^{t} \mathrm{d}s \, \frac{1}{4T} \left[\dot{x} + \frac{2}{\beta}\frac{\partial \phi}{\partial x}(x;\sigma) - \frac{\partial U}{\partial x}(x;\alpha) + f \right]^2. \quad (79)$$

Simple algebra shows that

$$\mathcal{I}_\mathrm{eff}[\mathbf{x},\sigma] = \mathcal{I}^\dagger[\mathbf{x};\sigma] - \int_{t'}^{t} \mathrm{d}s \left[\frac{\partial \phi}{\partial x}(x;\sigma) \left(\frac{\partial U}{\partial x}(x;\alpha) - f \right) - T \left(\frac{\partial \phi}{\partial x}(x;\sigma) \right)^2 \right]. \quad (80)$$

We know that, by definition, $\rho_\mathrm{SS}(x,\sigma)$ satisfies at fixed σ the following stationary Fokker–Planck equation:

$$T \frac{\partial^2 \rho_\mathrm{SS}}{\partial x^2}(x;\sigma) + \frac{\partial}{\partial x} \left[\left(\frac{\partial U}{\partial x}(x;\alpha) - f \right) \rho_\mathrm{SS}(x;\sigma) \right] = 0. \quad (81)$$

Now, putting $\rho_\mathrm{SS}(x;\sigma) = \exp[-\phi(x;\sigma)]$ in (81) one obtains

$$\left(\frac{\partial \phi}{\partial x} \right) \left(\frac{\partial U}{\partial x} - f \right) - T \left(\frac{\partial \phi}{\partial x} \right)^2 = \frac{\partial^2}{\partial x^2}(U - T\phi). \quad (82)$$

We thus get

$$\mathcal{I}_\mathrm{eff}[\mathbf{x};\sigma] = \mathcal{I}^\dagger[\mathbf{x};\sigma] + \frac{1}{2} \int_{t'}^{t} \mathrm{d}s \, \frac{\partial^2}{\partial x^2} \left[2T\phi(x;\sigma) - 2U(x;\alpha) \right]. \quad (83)$$

Substituting (83) in (74) yields

$$K^\dagger(x,t|x',t') = \Pi^{-1}(t-t') \int_{x(t')=x'}^{x(t)=x} Dx \, J_\mathrm{eff}[\mathbf{x};\sigma] \exp[-\mathcal{I}_\mathrm{eff}[\mathbf{x};\sigma]], \quad (84)$$

where

$$J_{\rm eff}[\mathbf{x};\sigma] = J[\mathbf{x};\sigma]\exp\left\{\frac{1}{2}\int_{t'}^{t}{\rm d}s\,\frac{\partial^2}{\partial x^2}\left[\frac{2}{\beta}\phi(x;\sigma)-2U(x;\alpha)\right]\right\}$$
$$= \exp\left\{\frac{1}{2}\int_{t'}^{t}{\rm d}s\,\frac{\partial^2}{\partial x^2}\left[\frac{2}{\beta}\phi(x;\sigma)-U(x;\alpha)\right]\right\}. \quad (85)$$

Let us define now the effective potential:

$$V_{\rm eff}(x;\sigma) = \frac{2}{\beta}\phi(x;\sigma) - U(x;\alpha). \quad (86)$$

With this definition we have that

$$\mathcal{I}_{\rm eff}[\mathbf{x};\sigma] = \left[\dot{x} + \frac{\partial}{\partial x}V_{\rm eff}(x;\sigma) + f\right]^2, \qquad J_{\rm eff}[\mathbf{x};\sigma] = \exp\left[\frac{1}{2}\int_{t'}^{t}{\rm d}s\,\frac{\partial^2}{\partial x^2}V_{\rm eff}(x;\sigma)\right]. \quad (87)$$

From (84) and (87) one obtains that the dual propagator is the propagator corresponding, in the Stratonovich scheme, to the Langevin equation:

$$\dot{x} = -\frac{\partial}{\partial x}V_{\rm eff}(x;\sigma) - f + \xi, \quad (88)$$

where, as can be seen from (56), the force has reversed sign. It is easy now to see that, when detailed balance holds, this dynamics coincides with the real dynamics (56). In fact, in this case we have $\phi(x;\sigma) = (1/T)[U(x) - fx - F(T;\sigma)]$, where $F(T;\sigma)$ is the free energy. Substitution of this choice for ϕ in (88) directly leads to (56). Concluding this discussion, we remark that the dual dynamics corresponds to the dynamics of a system having the same steady-state probability density function but with opposite probability current in the steady state.

Finally we can also compute $\mathcal{S}_{\rm na}$ directly from the path-integral representation of the Langevin dynamics as

$$\mathcal{S}_{\rm na} = I_{\rm R}^{\dagger}[\mathbf{x}^{\rm R};\sigma^{\rm R}] - I[\mathbf{x};\sigma] = \ln\frac{p_0(x(t');\sigma(t'))}{p_1(x(t);\sigma(t))} + \mathcal{I}^{\dagger}[\mathbf{x}^{\rm R};\sigma^{\rm R}] - \mathcal{I}[\mathbf{x};\sigma]$$
$$= \ln\frac{p_0(x(t');\sigma(t'))}{p_1(x(t);\sigma(t))} - \int_{t'}^{t}{\rm d}s\,\frac{\partial\phi}{\partial x}(x;\sigma)\dot{x}. \quad (89)$$

Taking now $t'=0$ and $t=\tau$, we obtain (23).

3.1.3. Derivation of known FTs. We now use the general symmetry equation (45) and the two relevant decompositions obtained above for the trajectory entropy production in the considered case in order to reobtain, from a unified point of view, the known FTs. First one sees immediately that

$$P(S_{\rm s},S_{\rm r})/P^{\rm R}(-S_{\rm s},-S_{\rm r}) = {\rm e}^{S_{\rm s}+S_{\rm r}} \qquad \text{and} \qquad P(S_{\rm a},S_{\rm na})/P^{\rm R}(-S_{\rm a},-S_{\rm na}) = {\rm e}^{S_{\rm a}+S_{\rm na}}, \quad (90)$$

or equivalently

$$\langle {\rm e}^{-\lambda_1\mathcal{S}_{\rm s}-\lambda_2\mathcal{S}_{\rm r}}\rangle = \langle {\rm e}^{-(1-\lambda_1)\mathcal{S}_{\rm s}^{\rm R}-(1-\lambda_2)\mathcal{S}_{\rm r}^{\rm R}}\rangle_{\rm R} \qquad \text{and}$$
$$\langle {\rm e}^{-\lambda_1\mathcal{S}_{\rm a}-\lambda_2\mathcal{S}_{\rm na}}\rangle = \langle {\rm e}^{-(1-\lambda_1)\mathcal{S}_{\rm a}^{\rm R}-(1-\lambda_2)\mathcal{S}_{\rm na}^{\rm R}}\rangle_{\rm R}. \quad (91)$$

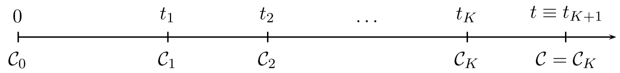

Figure 1. A history of configurations $\mathcal{C}_0 \to \cdots \to \mathcal{C}_K$. Between t_k and t_{k+1}, the system remains in configuration $\mathcal{C}_k$.

It is worth noting that these relations do not involve dual probability distribution functions (PDFs), and thus they can be tested for a physical system with a given dynamics. We also note that, while one can show both that $S_{\rm a}$ and $S_{\rm na}$ satisfy separately a DFT by using dual PDFs [19], $S_{\rm s}$ and $S_{\rm r}$ satisfy a joint DFT although they do not satisfy separately a DFT.

Let us now derive from a unified view the known FTs. Using equation (77) we can introduce the dual joint probability density function:

$$P^\dagger(S_{\rm a}, S_{\rm na}) = P(-S_{\rm a}, S_{\rm na})\mathrm{e}^{S_{\rm a}}. \tag{92}$$

This probability density function is, by virtue of equation (78), correctly normalized:

$$\begin{aligned}\int \mathrm{d}S_{\rm a}\,\mathrm{d}S_{\rm na} P^\dagger(S_{\rm a}, S_{\rm na}) &= \int \mathrm{d}S_{\rm a}\,\mathrm{d}S_{\rm na} P(-S_{\rm a}, S_{\rm na})\mathrm{e}^{S_{\rm a}} = \\ &= \int \mathrm{d}S_{\rm a}\,\mathrm{d}S_{\rm na} P(S_{\rm a}, S_{\rm na})\mathrm{e}^{-S_{\rm a}} = \langle \mathrm{e}^{-S_{\rm a}}\rangle = 1.\end{aligned} \tag{93}$$

Let us first derive the Speck–Seifert DFT [10]. We have

$$P(S_{\rm a}) = \int \mathrm{d}S_{\rm na}\, P(S_{\rm a}, S_{\rm na}) = \int \mathrm{d}S_{\rm na}\, P^\dagger(-S_{\rm a}, S_{\rm na})\mathrm{e}^{S_{\rm a}} = P^\dagger(-S_{\rm a})\mathrm{e}^{S_{\rm a}}. \tag{94}$$

We can also derive the Chernyak–Chertkov–Jarzynski DFT [20] for the non-adiabatic contribution:

$$\begin{aligned}P(S_{\rm na}) &= \int \mathrm{d}S_{\rm a}\, P(S_{\rm a}, S_{\rm na}) = \int \mathrm{d}S_{\rm a}\, P^{\rm R}(-S_{\rm a}, -S_{\rm na})\mathrm{e}^{S_{\rm a}+S_{\rm na}} \\ &= \int \mathrm{d}S_{\rm a}\, P^{\dagger R}(S_{\rm a}, -S_{\rm na})\mathrm{e}^{S_{\rm na}} = P^{\dagger R}(-S_{\rm na})\mathrm{e}^{S_{\rm na}}.\end{aligned} \tag{95}$$

3.2. Markov dynamics

3.2.1. Settings. The symmetry (45) holds for any decomposition of the entropy as a sum of terms which are antisymmetric by time-reversal. The difficult step is to explicitly write such kinds of decompositions. In this section, we consider a continuous-time Markov process for a system described by discrete configurations $\{\mathcal{C}\}$ and work out such decompositions for that dynamics. The master equation for the probability $P(\mathcal{C}, t)$ for the system to be in configuration $\mathcal{C}$ at time t is

$$\partial_t P(\mathcal{C}, t) = \sum_{\mathcal{C}'}[W(\mathcal{C}' \to \mathcal{C}, \sigma)P(\mathcal{C}', t) - W(\mathcal{C} \to \mathcal{C}', \sigma)P(\mathcal{C}, t)] \tag{96}$$

where σ is an external time-dependent control parameter. A system history consists of a sequence $\mathcal{C}_0, \ldots, \mathcal{C}_K$ of configurations with jumps at times $t_1, \ldots, t_K$ (see figure 1).

J. Stat. Mech. (2012) P02009

Let us introduce the dual transition rates $W^\dagger$:

$$W^\dagger(\mathcal{C}\to\mathcal{C}',\sigma)\equiv\frac{P_{\rm st}(\mathcal{C}',\sigma)}{P_{\rm st}(\mathcal{C},\sigma)}W(\mathcal{C}'\to\mathcal{C},\sigma)={\rm e}^{-[\phi(\mathcal{C}',\sigma)-\phi(\mathcal{C},\sigma)]}W(\mathcal{C}'\to\mathcal{C},\sigma). \tag{97}$$

Note that when detailed balance is obeyed (i.e. $P_{\rm eq}(\mathcal{C})W(\mathcal{C}\to\mathcal{C}')=P_{\rm eq}(\mathcal{C}')W(\mathcal{C}'\to\mathcal{C})$ for some equilibrium distribution $P_{\rm eq}(\mathcal{C})$), the dual dynamics is the same as the original one. We first remark that the dual escape rate $r^\dagger(\mathcal{C},\sigma)\equiv\sum_{\mathcal{C}'}W^\dagger(\mathcal{C}\to\mathcal{C}',\sigma)$ is the same as the original one:

$$r^\dagger(\mathcal{C},\sigma)=\frac{1}{P_{\rm st}(\mathcal{C},\sigma)}\sum_{\mathcal{C}'}W(\mathcal{C}'\to\mathcal{C},\sigma)P_{\rm st}(\mathcal{C}',\sigma)=\frac{1}{P_{\rm st}(\mathcal{C},\sigma)}r(\mathcal{C},\sigma)P_{\rm st}(\mathcal{C},\sigma)=r(\mathcal{C},\sigma). \tag{98}$$

More importantly, the *steady state* of the dual dynamics is the same since

$$\begin{aligned}\sum_{\mathcal{C}'}W^\dagger(\mathcal{C}'\to\mathcal{C},\sigma)P_{\rm st}(\mathcal{C}',\sigma)&=\sum_{\mathcal{C}'}W(\mathcal{C}\to\mathcal{C}',\sigma)P_{\rm st}(\mathcal{C},\sigma)\\&=r(\mathcal{C},\sigma)P_{\rm st}(\mathcal{C},\sigma)=r^\dagger(\mathcal{C},\sigma)P_{\rm st}(\mathcal{C},\sigma).\end{aligned} \tag{99}$$

3.2.2. Entropies. The (density of) probability of a trajectory specified as in figure 1 is

$${\rm Prob}[\mathcal{C},\sigma]={\rm e}^{-\int_0^\tau{\rm d}t\,r(\mathcal{C}(t),\sigma(t))}\prod_{k=1}^{K}W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})\;P_0(\mathcal{C}(0),\sigma(0)) \tag{100}$$

where K is the number of events and P_0 the initial distribution. It precisely means that the average of an observable $\mathcal{O}$ depending on the history of configurations and on the protocol σ is

$$\langle\mathcal{O}\rangle=\sum_{K\geq0}\sum_{\mathcal{C}_0\cdots\mathcal{C}_K}\int_0^t{\rm d}t_K\int_0^{t_K}{\rm d}t_{K-1}\ldots\int_0^{t_2}{\rm d}t_1\,\mathcal{O}[\mathcal{C},\sigma]\,{\rm Prob}[\mathcal{C},\sigma]. \tag{101}$$

By analogy to systems described by Langevin dynamics, one defines the total entropy:

$$\mathcal{S}[\mathcal{C},\sigma]=\log\frac{{\rm Prob}[\mathcal{C},\sigma]}{{\rm Prob}^{\rm R}[\mathcal{C}^{\rm R},\sigma^{\rm R}]} \tag{102}$$

$$=\sum_{k=1}^{K}\log\frac{W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})}{W(\mathcal{C}_k\to\mathcal{C}_{k-1},\sigma_{t_k})}+\log\frac{P_0(\mathcal{C}(0),\sigma(0))}{P_1(\mathcal{C}(t),\sigma(t))} \tag{103}$$

where $\mathcal{C}^{\rm R}$ is the reversed trajectory and $\sigma^{\rm R}$ the reverse protocol. Note that $\mathcal{S}[\mathcal{C},\sigma]=-\mathcal{S}[\mathcal{C}^{\rm R},\sigma^{\rm R}]$. We now would like to split the action into a sum of different terms, each of them also *antisymmetric* upon time-reversal. To do so, we assume that $P_0=P_1$ is the steady state $P_{\rm st}={\rm e}^{-\phi}$. Then, we define the 'house-keeping' entropy $\mathcal{Q}^{\rm hk}$ as

$$\beta\mathcal{Q}^{\rm hk}[\mathcal{C},\sigma]=\sum_{k=1}^{K}\log\frac{W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})}{W^\dagger(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})} \tag{104}$$

$$=\sum_{k=1}^{K}\log\frac{W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})}{W(\mathcal{C}_k\to\mathcal{C}_{k-1},\sigma_{t_k})}+\sum_{k=1}^{K}\phi(\mathcal{C}_k,\sigma_{t_k})-\phi(\mathcal{C}_{k-1},\sigma_{t_k}), \tag{105}$$

J. Stat. Mech. (2012) P02009

Note that it is indeed antisymmetric upon time-reversal: $\beta\mathcal{Q}^{\rm hk}[\mathcal{C},\sigma] = -\beta\mathcal{Q}^{\rm hk}[\mathcal{C}^{\rm R},\sigma^{\rm R}]$. Moreover, we see by direct computation (e.g. from (112) below) that the total entropy $\mathcal{S}$ is written in terms of the house-keeping work as

$$\mathcal{S}[\mathcal{C},\sigma] = \beta\mathcal{Q}^{\rm hk}[\mathcal{C},\sigma] + \mathcal{Y}[\mathcal{C},\sigma] \tag{106}$$

where

$$\mathcal{Y}[\mathcal{C},\sigma] = \int_0^\tau {\rm d}t\,\dot\sigma\frac{\partial\phi}{\partial\sigma} \tag{107}$$

is the Hatano–Sasa functional. Note that each term of the decomposition is antisymmetric.

Note finally that defining $\Delta\phi = \phi(\mathcal{C}(t),\sigma(t)) - \phi(\mathcal{C}(0),\sigma(0))$ and

$$\beta\mathcal{Q}^{\rm tot}[\mathcal{C},\sigma] = \sum_{k=1}^K \log\frac{W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})}{W(\mathcal{C}_k\to\mathcal{C}_{k-1},\sigma_{t_k})} \tag{108}$$

one reads directly from the definition (103) that

$$\mathcal{S}[\mathcal{C},\sigma] = \beta\mathcal{Q}^{\rm tot}[\mathcal{C},\sigma] + \Delta\phi \tag{109}$$

where again each term of this decomposition is antisymmetric.

To summarize:

$$\begin{aligned}\mathcal{S}[\mathcal{C},\sigma] = {} & \overbrace{\beta\mathcal{Q}^{\rm tot}[\mathcal{C},\sigma]}^{\mathcal{S}_1} + \overbrace{\sum_{k=1}^K[\phi(\mathcal{C}_k,\sigma_{t_k}) - \phi(\mathcal{C}_{k-1},\sigma_{t_k})]}^{\mathcal{S}_2} \\ & - \overbrace{\sum_{k=1}^K[\phi(\mathcal{C}_k,\sigma_{t_k}) - \phi(\mathcal{C}_{k-1},\sigma_{t_k})]}^{\mathcal{S}_2} + \overbrace{\Delta\phi}^{\mathcal{S}_3}\ .\end{aligned} \tag{110}$$

The first decomposition consists in grouping $\beta\mathcal{Q}^{\rm hk}[\mathcal{C},\sigma] = \mathcal{S}_1 + \mathcal{S}_2$ and $\mathcal{Y}[\mathcal{C},\sigma] = -\mathcal{S}_2 + \mathcal{S}_3$ while the second decomposition simply corresponds to $\beta\mathcal{Q}^{\rm tot}[\mathcal{C},\sigma] = \mathcal{S}_1$ and $\Delta\phi = \mathcal{S}_3$.

We also remark that writing

$$\mathcal{S}[\mathcal{C},\sigma] = \beta\mathcal{Q}^{\rm hk}[\mathcal{C},\sigma] - \sum_{k=1}^K[\phi(\mathcal{C}_k,\sigma_{t_k}) - \phi(\mathcal{C}_{k-1},\sigma_{t_k})] + \Delta\phi \tag{111}$$

one has a decomposition into a sum of *three* antisymmetric terms implying FTs using (45).

3.2.3. Link to the Hatano–Sasa functional and symmetries of operators. It is instructive to rewrite the Hatano–Sasa functional $\mathcal{Y}$ as

$$\begin{aligned}\mathcal{Y}[\mathcal{C},\sigma] &= \int_0^\tau {\rm d}t\,\dot\sigma\frac{\partial\phi}{\partial\sigma} = \sum_{k=0}^K\int_{t_k}^{t_{k+1}} {\rm d}t\,\dot\sigma\frac{\partial\phi}{\partial\sigma} = \sum_{k=0}^K \phi(\mathcal{C}_k,\sigma_{t_{k+1}}) - \phi(\mathcal{C}_k,\sigma_{t_k}) \\ &= [\phi(\mathcal{C},\sigma)]_0^\tau - \sum_{k=1}^K \phi(\mathcal{C}_k,\sigma_{t_k}) - \phi(\mathcal{C}_{k-1},\sigma_{t_k}).\end{aligned} \tag{112}$$

J. Stat. Mech. (2012) P02009

We thus have split $\mathcal{Y}$ in two parts, $[\phi(\mathcal{C},\sigma)]_0^\tau$, which depends explicitly on the final time τ, and the reduced Hatano–Sasa functional:

$$\hat{h}_\tau = -\sum_{k=1}^{K} \phi(\mathcal{C}_k, \sigma_{t_k}) - \phi(\mathcal{C}_{k-1}, \sigma_{t_k}) \tag{113}$$

which (for fixed protocol $\sigma(t)$) depends only on the sequence of visited configurations $\mathcal{C}_k$ and of the jump times t_k. This decomposition helps us to write $\langle \mathrm{e}^{-s\mathcal{Y}}\rangle$ in terms of an s-modified evolution operator. We show in appendix A that the DFTs arising from the decomposition (110) of the dynamical entropy translate into symmetries of a modified operator of evolutions $\mathbb{W}_\lambda$ such that $\langle \mathrm{e}^{-\lambda S}\rangle = \langle P_1|\mathcal{T}\mathrm{e}^{t\mathbb{W}_\lambda}|P_0\rangle$, where $\mathcal{T}\mathrm{e}^{\cdot}$ is the time-ordered exponential. Compared to the original approach of Hatano and Sasa [9] our derivations bring only into play the properties of the operator of evolution (through the use of the time-ordered exponential) and do not require a time discretization.

4. Applications

4.1. Experimental errors

4.1.1. General derivation. We now show with some detail how joint FTs provide insights into the experimental error in the evaluation of entropy productions. The results we have shown up to now are rather general and do not depend on the shape of the initial and final PDFs for the considered systems under arbitrary protocols. For concreteness and in order to be closer to experiments, we consider in this section the case of transitions between NESS, so the initial and final PDFs for the system are taken as the ones corresponding to the steady states of the system with the given values of the parameters.

In this situation, the non-adiabatic contribution to the total trajectory entropy production is simply the Hatano–Sasa (HS) functional [9]:

$$\mathcal{Y}[\mathbf{x};\sigma] = \int_0^\tau \mathrm{d}t\, \dot{\sigma}\frac{\partial\phi}{\partial\sigma}(x;\sigma). \tag{114}$$

We are interested in experimental measurements of this quantity, performed for instance in [31]. In a typical experiment one follows in general the following protocol:

- First, one measures the particle density in different points and build an histogram from where the steady-state PDF of the system can be inferred for different values of the control parameters (and correspondingly $\phi_{\mathrm{exp}}(x;\sigma) = -\ln\rho_{\mathrm{ss}}^{\mathrm{exp}}(x;\sigma)$).
- Second, one follows a given particle during a prescribed protocol (sampling its position) and builds, for the considered realization of the experiment, the Hatano–Sasa functional: $\mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}};\sigma] = \int_0^\tau \mathrm{d}t\,\dot{\sigma}(\partial\phi_{\mathrm{exp}}/\partial\sigma)(x_{\mathrm{exp}};\sigma)$.
- Finally, one averages over many experimental realizations the quantity of interest, for example one determines the average $\langle\exp[-\mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}};\sigma]]\rangle$ in order to test the validity of the Hatano–Sasa IFT.

Although these settings are very generic, there are cases where the steady-state PDF of the system is known and one do not need to consider this source of experimental errors, as in [31].

Let us first note that ϕ_{exp} and x_{exp} are not the true values for the quantities one is trying to determine in the experiment, but are outcomes from measurements. They thus implicitly carry errors. We can, in principle, repeat the experiment many times and perform a histogram in order to calculate the probability of the experimental values of the HS functional. Imagine the hypothetical situation in which we are also able to directly obtain the true value of the HS functional, from where we can extract for each experiment the corresponding deviation of the measured value of $\mathcal{Y}$ with respect to the real one: $\mathcal{E}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}}, x; \sigma] = \mathcal{Y}[\mathbf{x}; \sigma] - \mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}}; \sigma]$. In that case we are able to experimentally build the joint PDF:

$$P(Y_{\mathrm{exp}}, E_{\mathrm{exp}}) = \int_{-i\infty}^{+i\infty} \frac{\mathrm{d}\lambda_1\, \mathrm{d}\lambda_2}{(2\pi)^2} \mathrm{e}^{\lambda_1 Y_{\mathrm{exp}} + \lambda_2 E_{\mathrm{exp}}} G_{\mathrm{exp}}(\lambda_1, \lambda_2), \tag{115}$$

$$G_{\mathrm{exp}}(\lambda_1, \lambda_2) = \overline{\langle \mathrm{e}^{-\lambda_1 \mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}};\sigma] - \lambda_2 \mathcal{E}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}}, x;\sigma]} \rangle}. \tag{116}$$

In equation (116), the brackets denote the thermal average while the overbar denotes the average over the distribution of experimental measurement errors. Let the conditional probability of position outcomes be $\mathcal{P}_{\mathrm{err}}[\mathbf{x}_{\mathrm{exp}}|x]$, that is, the probability of obtaining the trajectory of outcomes $x_{\mathrm{exp}}(t)$, $x(t)$ being the true trajectory of the particle. In this case, we can write a mathematical expression for the experimental generating function (GF):

$$G_{\mathrm{exp}}(\lambda_1, \lambda_2) = \int \mathcal{D}x \mathcal{D}x_{\mathrm{exp}} \mathcal{P}[\mathbf{x}; \sigma] \mathcal{P}_{\mathrm{err}}[\mathbf{x}_{\mathrm{exp}}|x] \exp\{-\lambda_1 \mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}}; \sigma] - \lambda_2 \mathcal{E}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}}, x; \sigma]\}. \tag{117}$$

Note that, by a suitable rearrangement of terms, we can rewrite the previous expressions as

$$G_{\mathrm{exp}}(\lambda_1, \lambda_2) = \int \mathcal{D}x \mathcal{P}[\mathbf{x}; \sigma] \mathrm{e}^{-\lambda_2 \mathcal{Y}[\mathbf{x};\sigma]} \int \mathcal{D}x_{\mathrm{exp}} \mathcal{P}_{\mathrm{err}}[\mathbf{x}_{\mathrm{exp}}|x] \mathrm{e}^{-(\lambda_1 - \lambda_2)\mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}};\sigma]}. \tag{118}$$

Let us introduce the quantity

$$\mathcal{B}[\lambda_1, \lambda_2, x; \sigma] = \int \mathcal{D}x_{\mathrm{exp}} \mathcal{P}_{\mathrm{err}}[\mathbf{x}_{\mathrm{exp}}|x] \mathrm{e}^{-(\lambda_1 - \lambda_2)\mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}};\sigma]}. \tag{119}$$

Then equation (118) can be written as follows:

$$G_{\mathrm{exp}}(\lambda_1, \lambda_2) = \langle \mathcal{B}[\lambda_1, \lambda_2, x; \sigma] \mathrm{e}^{-\lambda_2 \mathcal{Y}[\mathbf{x};\sigma]} \rangle = \langle \mathcal{B}[\lambda_1, \lambda_2, x_{\mathrm{R}}; \sigma] \mathrm{e}^{-\lambda_2 \mathcal{Y}[\mathbf{x}_{\mathrm{R}};\sigma] - \mathcal{S}[\mathbf{x};\sigma_{\mathrm{R}}]} \rangle_{\mathrm{R}}, \tag{120}$$

where the second equality follows from our generalized Crooks-like relation (see section 2.1), identifying the observable $\mathcal{O}[\mathbf{x}; \sigma]$ with the quantity $\mathcal{B}[\lambda_1, \lambda_2, x; \sigma]\mathrm{e}^{-\lambda_2 \mathcal{Y}[\mathbf{x};\sigma]}$.

Now recalling that $\mathcal{Y}[\mathbf{x}_{\mathrm{R}}; \sigma] = -\mathcal{Y}[\mathbf{x}; \sigma_{\mathrm{R}}]$, and also that the total trajectory entropy production splits into the sum of a non-adiabatic contribution (in this case the Hatano–Sasa functional) and an adiabatic one, and that weighted averages with the exponential of the adiabatic contribution are equivalent to averages over trajectories given by the dual dynamics, one writes

$$G_{\mathrm{exp}}(\lambda_1, \lambda_2) = \langle \mathcal{B}[\lambda_1, \lambda_2, x_{\mathrm{R}}; \sigma] \mathrm{e}^{-(1-\lambda_2)\mathcal{Y}[\mathbf{x};\sigma_{\mathrm{R}}]} \rangle_{\mathrm{R}}^{\dagger}. \tag{121}$$

Let us now analyse the behaviour of the quantity $\mathcal{B}[\lambda_1, \lambda_2, x; \sigma]$, when a time-reversal operation is performed. For that purpose, let us impose some physically acceptable and

J. Stat. Mech. (2012) P02009

general properties on the conditional error distribution: we assume that the source of experimental errors associated with $\mathcal{P}[\mathbf{x}_{\mathtt{err}}|x]$ (the shape of the distribution) does not depend explicitly on time neither on the form of the experimental protocol. This error is basically related to some inherent properties of the measurement apparatus. On the other hand, we assume that there is no memory in this distribution: the outcome of the measurement at time t is only related to the true value of x at the same time (this has indeed being implicitly assumed in the notation for this probability density). We finally assume that this distribution does not depend on any time derivatives of the involved variables. From all these requirements, the distribution $\mathcal{P}[\mathbf{x}_{\mathtt{err}}|x]$ is invariant upon time-reversal:

$$\mathcal{P}[\mathbf{x}_{\mathtt{err}}|x] = \mathcal{P}[\mathbf{x}^{\mathrm{R}}_{\mathtt{err}}|x_{\mathrm{R}}]. \tag{122}$$

We can then write

$$\mathcal{B}[\lambda_1, \lambda_2, x_{\mathrm{R}}; \sigma] = \int \mathcal{D}x_{\mathtt{exp}} \mathcal{P}_{\mathtt{err}}[\mathbf{x}_{\mathtt{exp}}|x_{\mathrm{R}}] \mathrm{e}^{-(\lambda_1-\lambda_2)\mathcal{Y}_{\mathtt{exp}}[\mathbf{x}_{\mathtt{exp}};\sigma]}. \tag{123}$$

Make now the change of variables $x_{\mathtt{exp}} = z_{\mathrm{R}}$. Note that this transformation has a Jacobian identically equal to one and that $\mathcal{Y}_{\mathtt{exp}}[\mathbf{x}^{\mathrm{R}}_{\mathtt{exp}}, \sigma] = -\mathcal{Y}_{\mathtt{exp}}[\mathbf{x}_{\mathtt{exp}}, \sigma_{\mathrm{R}}]$, which, in fact, comes from the definition of $\mathcal{Y}_{\mathtt{exp}}$. We then have

$$\begin{aligned}\mathcal{B}[\lambda_1, \lambda_2, x_{\mathrm{R}}; \sigma] &= \int \mathcal{D}z_{\mathrm{R}} \mathcal{P}_{\mathtt{err}}[z_{\mathrm{R}}|x_{\mathrm{R}}] \mathrm{e}^{-(\lambda_1-\lambda_2)\mathcal{Y}_{\mathtt{exp}}[z_{\mathrm{R}};\sigma]} = \\ &= \int \mathcal{D}z \mathcal{P}_{\mathtt{err}}[z|x] \mathrm{e}^{-[(1-\lambda_1)-(1-\lambda_2)]\mathcal{Y}_{\mathtt{exp}}[z;\sigma_{\mathrm{R}}]},\end{aligned} \tag{124}$$

which in fact means

$$\mathcal{B}[\lambda_1, \lambda_2, x_{\mathrm{R}}; \sigma] = \mathcal{B}[1-\lambda_1, 1-\lambda_2, x; \sigma_{\mathrm{R}}]. \tag{125}$$

Using now (120), (121) and (125), we obtain

$$G_{\mathtt{exp}}(\lambda_1, \lambda_2) = G^{\dagger R}_{\mathtt{exp}}(1-\lambda_1, 1-\lambda_2), \tag{126}$$

from where one immediately obtains

$$P(Y_{\mathtt{exp}}, E_{\mathtt{exp}})/P^{\dagger R}(-Y_{\mathtt{exp}}, -E_{\mathtt{exp}}) = \mathrm{e}^{Y_{\mathtt{exp}}+E_{\mathtt{exp}}}. \tag{127}$$

From the last expression, or by simply using $\lambda_1 = 1$, $\lambda_2 = 0$ in (126) we obtain, for example, the relation

$$\overline{\langle \mathrm{e}^{-\mathcal{Y}_{\mathtt{exp}}[\mathbf{x}_{\mathtt{exp}};\sigma]} \rangle} = \overline{\langle \mathrm{e}^{-\mathcal{E}_{\mathtt{exp}}[\mathbf{x},x_{\mathtt{exp}};\sigma]} \rangle}^{\dagger}_{\mathrm{R}}. \tag{128}$$

From the analysis of $\overline{\langle \mathrm{e}^{-\mathcal{E}_{\mathtt{exp}}} \rangle}^{\dagger}_{\mathrm{R}}$ for specific cases of experimental errors, one can estimate the dispersion of the experimentally obtained $\overline{\langle \mathrm{e}^{-\mathcal{Y}_{\mathtt{exp}}} \rangle}$ around $\langle \mathrm{e}^{-\mathcal{Y}} \rangle = 1$.

J. Stat. Mech. (2012) P02009

4.1.2. Experimental errors for a driven particle in a harmonic trap. Let us now consider for concreteness an exactly solvable case illustrating the validity of the derivation we made above. We focus on the experiment of Trepagnier *et al* [31]. In the considered experimental situation, a microscopic bead is dragged through water by using a steerable harmonic optical trap. In this case, treating the bead as a Brownian particle in a harmonic potential, the steady-state distribution for the system was obtained theoretically. In the experiment the velocity of the trap plays the role of the tunable control parameter. For more details on the experimental set-up and on the definitions, see [31]. In the given case, the Hatano–Sasa functional is

$$\mathcal{Y}[\mathbf{x}; v] = \frac{\beta\gamma}{\kappa} \int_0^\tau \mathrm{d}t\, \dot{v}(\kappa x + \gamma v), \tag{129}$$

where κ is the trap constant, γ is the friction coefficient of the bead in the solution and β is the inverse temperature.

The experimental Hatano–Sasa functional and the error can then be written as follows:

$$\mathcal{Y}_{\mathtt{exp}}[\mathbf{x}_{\mathtt{exp}}; v] = \frac{\beta\gamma}{\kappa} \int_0^\tau \mathrm{d}t\, \dot{v}(\kappa x_{\mathtt{exp}} + \gamma v), \tag{130}$$

$$\mathcal{E}_{\mathtt{exp}}[\mathbf{x}, x_{\mathtt{exp}}; v] = \beta\gamma \int_0^\tau \mathrm{d}t\, \dot{v}(x - x_{\mathtt{exp}}). \tag{131}$$

We assume that the general conditions we assumed for $\mathcal{P}_{\mathtt{err}}[\mathbf{x}|x_{\mathtt{exp}}] \equiv \mathcal{P}_{\mathtt{err}}[\mathbf{x} - x_{\mathtt{exp}}]$ are valid here. Consider now the generating functional of this conditional probability, which we denote by $G^J_{\mathtt{err}}$:

$$G^J_{\mathtt{err}}[J] = \int \mathcal{D}\eta \mathcal{P}_{\mathtt{err}}[\eta] \mathrm{e}^{-\mathrm{i}\int_0^\tau \mathrm{d}t J(t)\eta(t)}. \tag{132}$$

Then, rearranging equation (117) one has

$$\begin{aligned} G_{\mathtt{exp}}(\lambda_1, \lambda_2) &= \int \mathcal{D}x \mathcal{P}[\mathbf{x}; \sigma] \mathrm{e}^{-\lambda_1 \mathcal{Y}[\mathbf{x}; v]} \int \mathcal{D}x_{\mathtt{exp}} \mathcal{P}_{\mathtt{err}}[\mathbf{x} - x_{\mathtt{exp}}] \mathrm{e}^{-(\lambda_2 - \lambda_1)\mathcal{E}_{\mathtt{exp}}[\mathbf{x} - x_{\mathtt{exp}}; \sigma]} \\ &= G^J_{\mathtt{err}}[-\mathrm{i}(\lambda_2 - \lambda_1)\beta\gamma\dot{v}] \langle \mathrm{e}^{-\lambda_1 \mathcal{Y}[\mathbf{x}; v]} \rangle. \end{aligned} \tag{133}$$

Taking into account that $\dot{v} = -\dot{v}_{\mathrm{R}}$, we have

$$\begin{aligned} G_{\mathtt{exp}}(\lambda_1, \lambda_2) &= G^J_{\mathtt{err}}[-\mathrm{i}((1 - \lambda_2) - (1 - \lambda_1))\beta\gamma\dot{v}_{\mathrm{R}}] \langle \mathrm{e}^{-(1-\lambda_1)\mathcal{Y}[\mathbf{x}; v_{\mathrm{R}}]} \rangle^\dagger \\ &= G^{\dagger R}_{\mathtt{exp}}(1 - \lambda_1, 1 - \lambda_2), \end{aligned} \tag{134}$$

which is the result we have obtained before.

In order to conclude with this section, let us exploit the solvability of this model in order to show that indeed there is a link between the apparatus precision and the experimentally relevant measured quantities. For example, in this case one has

$$\overline{\langle \mathrm{e}^{-\mathcal{Y}_{\mathtt{exp}}[\mathbf{x}_{\mathtt{exp}}; v]} \rangle} = G_{\mathtt{exp}}(1, 0) \equiv G^J_{\mathtt{err}}[\mathrm{i}\beta\gamma\dot{v}], \tag{135}$$

from where we see that the experimental mean value of the exponential of the Hatano–Sasa functional is solely and directly determined by the distribution of measurement errors in the position of the particle. For example, in the case of an extremely precise

J. Stat. Mech. (2012) P02009

measurement, one has $G^J_{\mathtt{err}} = 1 \Rightarrow \overline{\langle \mathrm{e}^{-\mathcal{Y}_{\mathrm{exp}}} \rangle} = \langle \mathrm{e}^{-\mathcal{Y}} \rangle \equiv 1$. Consider now the case in which the experimental error conditional distribution is Gaussian, i.e.

$$\mathcal{P}_{\mathtt{err}}[\eta] \sim \exp\left\{ -\frac{1}{2\Delta^2} \int_0^\tau \mathrm{d}t\, \eta^2(t) \right\} \Rightarrow G^J_{\mathtt{err}}[J] = \exp\left\{ -\frac{\Delta^2}{2} \int_0^\tau \mathrm{d}t\, J^2(t) \right\}. \tag{136}$$

In this case, one sees that the experimental deviation from the Hatano–Sasa IFT is linked to the accuracy of the measurement apparatus:

$$-\log \overline{\langle \mathrm{e}^{-\mathcal{Y}_{\mathrm{exp}}[\mathbf{x}_{\mathrm{exp}};v]} \rangle} = -\frac{\Delta^2 \beta^2 \gamma^2}{2} \int_0^\tau \mathrm{d}t\, \dot{v}^2(t) < 0, \tag{137}$$

which is compatible with the experimental results of [31]. We learn from the last expression that, at very high temperatures, the 'violation' of the fluctuation theorem introduced by the lack of accuracy of the apparatus can be cured by thermal fluctuations. On the other hand, rapidly varying protocols may induce a remarkable enhancement of the referred-to 'violation' factor.

4.2. Generalized FDT, identities on correlation functions

In this section, we use the second equality in (91) in order to find some symmetries on correlation functions related to a generalized fluctuation–dissipation theorem (FDT) derived, for example, in [32]–[39] (see [40] for a review). Let us consider the situation in which the system, initially prepared in an NESS, is submitted to the variation of its parameters $\sigma_i(t) = \sigma_i^0 + \delta\sigma_i(t)$ in such a way that $|\delta\sigma_i(t)/\sigma_i^0| \ll 1$, with $\sigma_i^0 = \sigma_i(0)$ and $\delta\sigma_i(0) = 0$. In this context, within the particular form of [36], this generalized FDT can be written as follows:

$$\left\langle \frac{\partial \phi}{\partial \sigma_i}(x(t), \sigma_0) \right\rangle = \sum_j \int_0^t \chi_{ij}(t - t')\delta\sigma_j(t')\, \mathrm{d}t', \tag{138}$$

where

$$\chi_{ij}(t - t') = \frac{\mathrm{d}}{\mathrm{d}t} C_{ij}(t - t') = \frac{\mathrm{d}}{\mathrm{d}t} \left\langle \frac{\partial \phi}{\partial \sigma_i}(x(t), \sigma_0) \frac{\partial \phi}{\partial \sigma_j}(x(t'), \sigma_0) \right\rangle_{\mathtt{ss}}. \tag{139}$$

For a system performing equilibrium dynamics, where detailed balance holds, this correlation function satisfies the well-known Onsager symmetry relation $C_{ij}(t - t') = C_{ij}(t' - t) \equiv C_{ji}(t - t')$. However, in the most general case of an NESS, this relation breaks down because of the lack of detailed balance. However, a similar relation can be easily obtained for this case: $C_{ij}(t - t') = C^\dagger_{ij}(t' - t)$. The physical meaning of this expressions is clear. One is performing a time-reversal operation. When detailed balance holds, there are no probability currents in the system and this implies invariance upon time-reversal. In this case the system and its dual are the same. When detailed balance breaks down, there is a finite current in the steady state, which is odd under time-reversal. In this case, the original dynamics is equivalent to the time-reversed dual one (which as expected has a current opposite to the original one). This relation, although conceptually clear, however, does not seems to be useful: it involves correlations in two different physical systems. The question which comes now is: does there exist a suitable correlation function capable to exhibit a symmetry involving only the system under study? The answer is yes. We now discuss briefly the physical grounds behind this correlation function.

J. Stat. Mech. (2012) P02009

4.2.1. Dynamical ensembles and weighted averages. In a full equivalence with equilibrium thermodynamic ensembles, one can build trajectory-based ensembles for stochastic non-equilibrium systems, together with quantities equivalent to the partition function and the free-energy—in a dynamical thermodynamic formalism [41]. For equilibrium systems with a Hamiltonian $H(C)$ depending on configurations in the phase space, one defines in the canonical ensemble a free energy $F(\beta) = -(1/\beta)\ln\sum_C \exp\{-\beta H(C)\}$ and averages of observables can be computed as $\langle y\rangle = \sum_C y(C)\exp\{\beta(F(\beta) - H(C))\}$. The inverse temperature β, is a Lagrange multiplier fixing the mean value of the energy, so that plugging a given β privileges in the sum the configurations compatible with a prescribed value of the mean energy.

Now consider, for a dynamical system, a time-extensive functional $\mathcal{K}[\mathbf{x}] = \int_0^t \mathrm{d}t'\, f(x(t'))$. The s ensemble, very similar to the canonical ensemble in equilibrium statistical mechanics, is defined by the average

$$\langle y\rangle^s = \frac{\int \mathcal{D}[\mathbf{x}]\mathcal{P}[\mathbf{x}]y[\mathbf{x}]\mathrm{e}^{-s\mathcal{K}[\mathbf{x}]}}{\int \mathcal{D}[\mathbf{x}]\mathcal{P}[\mathbf{x}]\mathrm{e}^{-s\mathcal{K}[\mathbf{x}]}}. \tag{140}$$

This approach has proven to be useful in the study of dynamical phase transitions [42]–[46].

The conceptual idea behind the use of the parameter s is the same as in the case of the inverse temperature in canonical ensembles: in the long time limit, fixing the value of s allows one to calculate averages of functionals, not for typical trajectories but for atypical ones, where the mean value of the time-extensive functional $\mathcal{K}[\mathbf{x}]$ is fixed. In other words, weighted averages select trajectories compatible with a prescribed physical scenario. The presence of this parameter imposes some constraints on the system in favour of rare events by introducing a well-chosen bias on the evolution of the system.

Even if the question of dealing experimentally with those biased trajectories is actually far from being answered in a closed form (see, however, [47]), the procedure is interesting from a conceptual and numerical [48]–[50] point of view.

Let us now turn again to our problem. The symmetric nature of the correlation function when the system is in equilibrium breaks down as detailed balance does for systems with finite currents in the steady state. However, one can guess that, if one finds an appropriate time-extensive functional defining weighted averages, the symmetry will be restored for a given value of the Lagrange multiplier. The only thing we need is to average, in non-equilibrium systems, only over trajectories compatible with some equilibrium effective dynamics. We deal with this in more detail in what follows.

4.2.2. Restoring the symmetry of the correlation function using weighted averages. We first combine equation (91) with equation (7) to write

$$\langle \mathcal{O}[\mathbf{x};\sigma]\mathrm{e}^{-\lambda_1\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]-\lambda_2\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma]}\rangle = \langle \mathcal{O}[\mathbf{x}^{\mathrm{R}};\sigma]\mathrm{e}^{-(1-\lambda_1)\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma^{\mathrm{R}}]-(1-\lambda_2)\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma^{\mathrm{R}}]}\rangle^{\mathrm{R}}. \tag{141}$$

Let us now choose $\mathcal{O}[\mathbf{x};\sigma] = \exp\{-\lambda_2\mathcal{F}[\mathbf{x};\sigma]\}$, with $\mathcal{F}[\mathbf{x};\sigma] = \mathcal{F}[\mathbf{x}^{\mathrm{R}};\sigma]$. With this we can write

$$\langle \mathrm{e}^{-\lambda_1\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]-\lambda_2\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma]-\lambda_2\mathcal{F}[\mathbf{x};\sigma]}\rangle = \langle \mathrm{e}^{-(1-\lambda_1)\mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma^{\mathrm{R}}]-(1-\lambda_2)\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma^{\mathrm{R}}]-\lambda_2\mathcal{F}[\mathbf{x};\sigma]}\rangle^{\mathrm{R}}. \tag{142}$$

Let us note now that we can write, to second order in the small perturbations $\delta\sigma$:

$$\begin{aligned} \mathrm{e}^{-\lambda_1 \mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma]} \approx 1 &- \lambda_1 \sum_i \int_0^\tau \mathrm{d}t\, \delta\dot\sigma_i(t) \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0) \\ &- \lambda_1 \sum_{ij} \int_0^\tau \mathrm{d}t\, \delta\dot\sigma_i(t)\delta\sigma_j(t) \frac{\partial^2 \phi}{\partial \sigma_i \partial \sigma_j}(x(t),\sigma_0) + \\ &+ \frac{\lambda_1^2}{2} \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t'\, \delta\dot\sigma_i(t)\delta\dot\sigma_j(t') \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0) \frac{\partial \phi}{\partial \sigma_j}(x(t'),\sigma_0), \end{aligned} \tag{143}$$

and

$$\begin{aligned} \mathrm{e}^{-(1-\lambda_1) \mathcal{S}_{\mathrm{na}}[\mathbf{x};\sigma^{\mathrm{R}}]} \approx 1 &- (1-\lambda_1) \sum_i \int_0^\tau \mathrm{d}t\, \delta\dot\sigma_i^{\mathrm{R}}(t) \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0) \\ &- (1-\lambda_1) \sum_{ij} \int_0^\tau \mathrm{d}t\, \delta\dot\sigma_i^{\mathrm{R}}(t)\delta\sigma_j^{\mathrm{R}}(t) \frac{\partial^2 \phi}{\partial \sigma_i \partial \sigma_j}(x(t),\sigma_0) \\ &+ \frac{(1-\lambda_1)^2}{2} \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t'\, \delta\dot\sigma_i^{\mathrm{R}}(t)\delta\dot\sigma_j^{\mathrm{R}}(t') \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0) \frac{\partial \phi}{\partial \sigma_j}(x(t'),\sigma_0). \end{aligned} \tag{144}$$

Plugging (143) and (144) in (142) and equating the terms with the same power of λ_1, we obtain a set of equalities. Let us consider the one corresponding to λ_1^2:

$$\begin{aligned} &\left\langle \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t' \delta\dot\sigma_i(t)\delta\dot\sigma_j(t') \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0) \frac{\partial \phi}{\partial \sigma_j}(x(t'),\sigma_0) \mathrm{e}^{-\lambda_2 \mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0] - \lambda_2 \mathcal{F}[\mathbf{x};\sigma_0]} \right\rangle_{\mathrm{ss}} \\ &\quad = \left\langle \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t'\, \delta\dot\sigma_i^{\mathrm{R}}(t)\delta\dot\sigma_j^{\mathrm{R}}(t') \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0) \frac{\partial \phi}{\partial \sigma_j}(x(t'),\sigma_0) \right. \\ &\qquad \left. \times\ \mathrm{e}^{-(1-\lambda_2) \mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0] - \lambda_2 \mathcal{F}[\mathbf{x};\sigma_0]} \right\rangle_{\mathrm{ss}}. \end{aligned} \tag{145}$$

In the previous equation, the averages and the functionals in the exponentials can be taken in the steady state $\sigma = \sigma_0$, because we are considering only the second order in $\delta\sigma$. In the second term of the previous equation, we perform the change of variables $t \to \tau - t$ and $t' \to \tau - t'$ in the double integral. Then, the last equation is equivalent to the following identity:

$$\begin{aligned} &\left\langle \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t' \delta\dot\sigma_i(t)\delta\dot\sigma_j(t') \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0) \frac{\partial \phi}{\partial \sigma_j}(x(t'),\sigma_0) \mathrm{e}^{-\lambda_2 \mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0] - \lambda_2 \mathcal{F}[\mathbf{x};\sigma_0]} \right\rangle_{\mathrm{ss}} \\ &\quad = \left\langle \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t'\, \delta\dot\sigma_i(t)\delta\dot\sigma_j(t') \frac{\partial \phi}{\partial \sigma_i}(x^{\mathrm{R}}(t),\sigma_0) \frac{\partial \phi}{\partial \sigma_j}(x^{\mathrm{R}}(t'),\sigma_0) \right. \\ &\qquad \left. \times\ \mathrm{e}^{-(1-\lambda_2) \mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0] - \lambda_2 \mathcal{F}[\mathbf{x};\sigma_0]} \right\rangle_{\mathrm{ss}}. \end{aligned} \tag{146}$$

J. Stat. Mech. (2012) P02009

If we finally take $\lambda_2 = 1/2$, we obtain a symmetric relation

$$\left\langle \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t' \delta\dot{\sigma}_i(t)\delta\dot{\sigma}_j(t') \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0)\frac{\partial \phi}{\partial \sigma_j}(x(t'),\sigma_0)\mathrm{e}^{-(1/2)(\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0]+\mathcal{F}[\mathbf{x};\sigma_0])} \right\rangle_{\mathrm{ss}}$$
$$= \left\langle \sum_{ij} \int_0^\tau \mathrm{d}t\, \mathrm{d}t'\, \delta\dot{\sigma}_i(t)\delta\dot{\sigma}_j(t') \frac{\partial \phi}{\partial \sigma_i}(x^{\mathrm{R}}(t),\sigma_0)\frac{\partial \phi}{\partial \sigma_j}(x^{\mathrm{R}}(t'),\sigma_0) \times\ \mathrm{e}^{-(1/2)(\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0]+\mathcal{F}[\mathbf{x};\sigma_0])} \right\rangle_{\mathrm{ss}} . \qquad (147)$$

From the last expression, we can write

$$\left\langle \mathcal{C}_{ij}(\tau,0)\mathrm{e}^{-(1/2)(\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0]+\mathcal{F}[\mathbf{x};\sigma_0])}\right\rangle_{\mathrm{ss}} = \left\langle \mathcal{C}_{ij}(0,\tau)\mathrm{e}^{-(1/2)(\mathcal{S}_{\mathrm{a}}[\mathbf{x};\sigma_0]+\mathcal{F}[\mathbf{x};\sigma_0])}\right\rangle_{\mathrm{ss}}, \qquad (148)$$

where

$$\mathcal{C}_{ij}(t,t') = \frac{\partial \phi}{\partial \sigma_i}(x(t),\sigma_0)\frac{\partial \phi}{\partial \sigma_j}(x(t'),\sigma_0). \qquad (149)$$

Up to now, what we have done is very general and the functional $\mathcal{F}$ can be any quantity being invariant upon time-reversal. In the discussion below, we consider some physically relevant choice of this functional.

4.2.3. Mapping to an effective system with equilibrium dynamics. Let us consider the following Langevin evolution:

$$\dot{x} = -\frac{1}{\beta}\frac{\partial \phi}{\partial x}(x;\sigma) + \xi. \qquad (150)$$

We know that $\exp\{-\phi\}$ is a well-behaved distribution: it is continuous and with continuous derivatives at least up to second order. It is also correctly normalized. On the other hand, this function corresponds to the stationary solution of the Fokker–Planck equation associated with the process (150). This is thus the steady state of this equation which, if ϕ is nonsingular in any point of the space (excluding the infinite), is unique. Interestingly, we remark that for the dynamics (150) there are no currents in the steady state, so that detailed balance holds and the system performs equilibrium dynamics. The steady state is the same for both the system (150) and the original system under study described by the evolution equation (56).

The effective action of any process has always a part being odd upon time-reversal, which is related to entropy production, and an invariant part, which is related to the activity, which for Langevin dynamics in a potential $\mathcal{V}$ and with additive white noise, is

$$R_{\mathcal{V}}(\tau) = \int_0^\tau \mathrm{d}t \left[\frac{\beta}{4}\left(\frac{\partial \mathcal{V}}{\partial x}\right)^2 - \frac{1}{2}\frac{\partial^2 \mathcal{V}}{\partial x^2}\right]. \qquad (151)$$

The quantity $\exp(-\beta\, \mathrm{d}t(\mathrm{d}R/\mathrm{d}t))$ is proportional to the probability that the system has stayed in its configuration between t and $t+\mathrm{d}t$ [51]: in other words, $\beta|\mathrm{d}R/\mathrm{d}t|$ is the rate at which the system escapes its configuration and is related to the activity (or 'traffic') [41], [52]–[54].

J. Stat. Mech. (2012) P02009

Let us investigate how far is our original system from the 'equilibrium' system (150). For that we are going to introduce the distance between them in the usual way:

$$\Xi[\mathbf{x};\sigma] = \ln\left(\mathcal{P}[\mathbf{x};\sigma]/\mathcal{P}_\phi[\mathbf{x};\sigma]\right) = \tfrac{1}{2}\left(\mathcal{S}_\mathrm{a}[\mathbf{x};\sigma] + 2\left(R_\phi[\mathbf{x};\sigma] - R_U[\mathbf{x};\sigma]\right)\right) \quad (152)$$

such that $\langle \Xi[\mathbf{x};\sigma]\rangle \geq 0$ is the Kullback–Leibler distance between $\mathcal{P}[\mathbf{x};\sigma]$ and $\mathcal{P}_\phi[\mathbf{x};\sigma]$. This quantity satisfies the corresponding IFT:

$$\langle \mathrm{e}^{-\Xi[\mathbf{x};\sigma]}\rangle = \langle 1\rangle_\phi = 1. \quad (153)$$

Considering then the s ensemble built using the quantity Ξ, we constrain our system to trajectories compatible with the dynamics (150). In particular, averages over the steady state of the original system lacking detailed balance are directly mapped to an equilibrium steady state, with zero currents. Defining this average as

$$\langle \mathcal{O}\rangle^s = \frac{\left\langle \mathcal{O}[\mathbf{x};\sigma]\mathrm{e}^{-s\Xi[\mathbf{x};\sigma]}\right\rangle}{\left\langle \mathrm{e}^{-s\Xi[\mathbf{x};\sigma]}\right\rangle}, \quad (154)$$

we see that averages in the steady state with $s = 1$ are mapped to averages in equilibrium, $\langle\cdots\rangle_{\mathsf{ss}}^{s=1} \equiv \langle\cdots\rangle_{\mathsf{eq}}$. If we take in equation (148), $\mathcal{F} = 2(R_\phi - R_U)$, this weighted correlation function acquires an interesting physical meaning. In order to clarify it a little bit, let us consider the generalized FDT given by (138) in the case where the underlying dynamics is given by (150). In order to avoid confusions, let us clarify the notation even if it will be the same already used. For averages under the original dynamics with varying protocol σ, we use the symbol $\langle\cdots\rangle$ whereas for averages under the original dynamics with constant $\sigma = \sigma_0$, we use the notation $\langle\cdots\rangle_{\mathsf{ss}}$. On the other hand, for averages under the dynamics given by equation (150) with varying σ we use $\langle\cdots\rangle_\phi$ whereas for averages under this dynamics for constant $\sigma = \sigma_0$, we use $\langle\cdots\rangle_{\mathsf{eq}}$. We will have

$$\left\langle \frac{\partial\phi}{\partial\sigma_i}(x(t),\sigma_0)\right\rangle_\phi = \sum_j \int_0^t \chi_{ij}^{\mathsf{eq}}(t-t')\delta\sigma_j(t')\mathrm{d}t', \quad (155)$$

where

$$\chi_{ij}^{\mathsf{eq}}(t-t') = \frac{\mathrm{d}}{\mathrm{d}t}C_{ij}^{\mathsf{eq}}(t-t') = \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\partial\phi}{\partial\sigma_i}(x(t),\sigma_0)\frac{\partial\phi}{\partial\sigma_j}(x(t'),\sigma_0)\right\rangle_{\mathsf{eq}}. \quad (156)$$

To the same order in $\delta\sigma$, this implies

$$\left\langle \frac{\partial\phi}{\partial\sigma_i}(x(t),\sigma_0)\mathrm{e}^{-\Xi_0(t)}\right\rangle = \sum_j \int_0^t \chi_{ij}^W(t-t')\delta\sigma_j(t')\,\mathrm{d}t', \quad (157)$$

where

$$\chi_{ij}^W(t-t') = \frac{\mathrm{d}}{\mathrm{d}t}C_{ij}^W(t-t') = \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\partial\phi}{\partial\sigma_i}(x(t),\sigma_0)\frac{\partial\phi}{\partial\sigma_j}(x(t'),\sigma_0)\mathrm{e}^{-\Xi_0(t)}\right\rangle_{\mathsf{ss}}, \quad (158)$$

and $\Xi_0(t) = \Xi[\mathbf{x};\sigma_0]$. Let us now introduce the notations $a_i(t) = (\partial\phi/\partial\sigma_i)(x(t),\sigma_0)$ and $b_i(t) = (\partial\phi/\partial\sigma_i)(x(t),\sigma_0)\mathrm{e}^{-\Xi_0(t)}$. Then, one can rewrite the two previous equations as

$$\langle b_i(t)\rangle = \sum_j \int_0^t \chi_{ij}^{ba}(t-t')\delta\sigma_j(t')\,\mathrm{d}t', \quad (159)$$

with

$$\chi_{ij}^{ba}(t-t') = \frac{\mathrm{d}}{\mathrm{d}t}C_{ij}^{ba}(t-t') = \frac{\mathrm{d}}{\mathrm{d}t}\langle b_i(t)a_j(t')\rangle_{\mathsf{ss}}. \quad (160)$$

Then, equation (148) corresponds to the symmetry upon indices interchange:

$$C_{ij}^{ba}(t-t') = C_{ji}^{ba}(t-t'). \quad (161)$$

5. Conclusion

We have reviewed, for continuum and discrete systems following a Markovian dynamics, how symmetries of the dynamical entropy translate into detailed or integrated fluctuation theorems, in terms of the symmetries under particular transformations of joint probability distributions.

From a conceptual point of view, the FTs given in (90) provide a convenient perspective to understand the interplay between different contributions to the total entropy production. The first of these equalities shows an explicit relation between the entropy production fluctuations associated with the system and its environment, valid at finite time and without quasi-static assumption. It allows one to study the correlations between the two sources of entropy production. The second relation unifies the three detailed FTs derived in [19]. The relevance of this expression is twofold. First, it implies the previously known FTs and, secondly, it gives a relation between the fluctuations of these two different entropy productions without relying on the dual dynamics (which is useful from an experimental point of view since one does not need to run a second system with the dual dynamics). In practice, to compute the contributions to the total entropy production, one needs to measure the ratios between transition rates, the PDF of the system at initial and final times, and the PDF in the steady state—which is achievable (see [55] where ratios between transitions rates have been measured). In the case of continuous Langevin equations, only the PDFs have to be measured. In any case, (90) is thus of experimental use.

It is known that fluctuation theorems provide a non-equilibrium way to measure physically relevant properties of single molecules (such as conformational free energy differences, see [2] for a review). We have proposed two other outcomes: a way to estimate experimental errors due to fluctuations in measurement apparatus and a modified non-equilibrium fluctuation–dissipation relation (FDR) (see also [36]–[39]).

In the first case, we have studied the experimental measurements of, for example, the Hatano–Sasa functional. We have demonstrated that the joint PDF of the experimentally measured values and the error of each measurement satisfies an FT, which allow us to extract precise information about the influence of error sources on the results. In particular, we have obtained an exact expression for the 'violation factor' of the Hatano–Sasa IFT in the simple case considered in the experiment of Trepagnier *et al* [31]. It is worth noting that our results are compatible with the experimental measurements. We have explicitly obtained that, for protocols with higher dissipation, the experimental results are further from the expected ones than those of protocols with lower dissipation. This fact is observed in the broadening of error bars as dissipation increases. Using the same ideas, one can study some other family of interesting systems: those with feedback, where the external protocol imposed on the system depends on the outcome of some measurement which can be carried out with some error.

In the second case, we have obtained new symmetries for correlation functions linked to modified FDRs. Those symmetries are based on the use of weighted averages which introduce some bias in the system such that, controlling this weight, one forces the system to prefer some trajectories in the phase space. We have indeed shown that, for a particular form of this weight, the dynamics of the system is exactly mapped to the dynamics of some equivalent equilibrium system, giving a physical meaning to these symmetries. Although the result looks artificial, it can be reformulated in terms of generalized forces and currents, giving the possibility to recover Onsager reciprocity in the linear regime in the vicinity of an NESS [56].

Finally, let us note that joint PDFs–FTs can be relevant in other contexts. Consider, for example, a set of strongly connected systems which do not satisfy an FT for any of the entropic contributions of its constituents. In this case, the most precise information one can obtain involves contributions from all the subsystems. This can be seen in classical systems but also in quantum situations. For example, some recent modified FTs have been obtained for counting statistics of electron transport in quantum dots, coupled to quantum point contacts which play the role of measurement apparatus continuously monitoring the system (see, for example, [57] for recent developments and a discussion on the effects of the back action of the environment on the system). An extension of our procedure to quantum systems may also give some general framework to tackle such problems[4].

Acknowledgments

This work was supported by CNEA, CONICET (PIP11220090100051) and ANPCYT (PICT2007886), VL was supported in part by the Swiss NSF under MaNEP and Division II and thanks CNEA for hospitality. ABK acknowledges Universidad de Barcelona, Ministerio de Ciencia e Innovación (Spain) and Generalitat de Catalunya for partial support through the I3 program.

Appendix. Symmetries of operators

In this appendix we show that the FTs are equivalent for continuous-time Markov processes to symmetries of the modified operator of evolution, in the spirit of Lebowitz and Spohn [6].

A.1. Operator approach

We take the notation of section 3.2. Let us denote $P(\mathcal{C}, \hat{h}, t)$ the probability of being in configuration $\mathcal{C}$ at time t, having observed a value $\hat{h}$ of the reduced Hatano–Sasa functional (113) and having started from configuration $\mathcal{C}_0$. The equation of evolution is

$$\partial_t P(\mathcal{C}, \hat{h}, t) = \sum_{\mathcal{C}'} [W(\mathcal{C}' \to \mathcal{C}, \sigma) P(\mathcal{C}', \hat{h} + \phi(\mathcal{C}, \sigma) - \phi(\mathcal{C}', \sigma), t) - W(\mathcal{C} \to \mathcal{C}', \sigma) P(\mathcal{C}, \hat{h}, t)]. \tag{A.1}$$

The Laplace transform $\hat{P}(\mathcal{C}, \lambda, t) = \int \mathrm{d}\hat{h}\, \mathrm{e}^{-\lambda \hat{h}} P(\mathcal{C}, \hat{h}, t)$ obeys

$$\partial_t \hat{P}(\mathcal{C}, \lambda, t) = \sum_{\mathcal{C}'} [\mathrm{e}^{-\lambda[\phi(\mathcal{C}', \sigma) - \phi(\mathcal{C}, \sigma)]} W(\mathcal{C}' \to \mathcal{C}, \sigma) \hat{P}(\mathcal{C}', \lambda, t)] \tag{A.2}$$

or, vectorially, $\partial_t |\hat{P}(\lambda, t)\rangle = \mathbb{W}_\lambda |\hat{P}(\lambda, t)\rangle$ with a time-dependent operator $\mathbb{W}_\lambda$ of elements

$$(\mathbb{W}_\lambda)_{\mathcal{C},\mathcal{C}'} = \mathrm{e}^{-\lambda[\phi(\mathcal{C}', \sigma) - \phi(\mathcal{C}, \sigma)]} W(\mathcal{C}' \to \mathcal{C}, \sigma) - \delta_{\mathcal{C},\mathcal{C}'} r(\mathcal{C}, \sigma) \tag{A.3}$$

[4] For systems where coherence is not important, described by an effective Pauli master equation, the mapping to classical systems is straightforward. Otherwise, we speculate that a Schwinger–Keldysh approach is relevant.

J. Stat. Mech. (2012) P02009

where $r(\mathcal{C},\sigma)$ is the escape rate from configuration $\mathcal{C}$, defined as

$$r(\mathcal{C},\sigma) \equiv \sum_{\mathcal{C}'} W(\mathcal{C}\to\mathcal{C}',\sigma). \tag{A.4}$$

Our quantity of interest, $\langle \mathrm{e}^{-\lambda\mathcal{Y}}\rangle \equiv \mathrm{e}^{\mu(\lambda,\tau;\sigma)}$ is

$$\langle \mathrm{e}^{-\lambda\mathcal{Y}}\rangle = \sum_{\mathcal{C}_0} P_{\mathrm{st}}(\mathcal{C}_0;\sigma_0)\int \mathrm{d}\hat{h}\sum_{\mathcal{C}} P(\mathcal{C},\hat{h},\tau)\mathrm{e}^{-\lambda\hat{h}}\mathrm{e}^{-\lambda[\phi(\mathcal{C},\sigma_\tau)-\phi(\mathcal{C}_0,\sigma_0)]} \tag{A.5}$$

$$= \sum_{\mathcal{C},\mathcal{C}_0} \mathrm{e}^{-\lambda\phi(\mathcal{C},\sigma_\tau)}\hat{P}(\mathcal{C},\lambda,\tau)\mathrm{e}^{-(1-\lambda)\phi(\mathcal{C}_0,\sigma_0)} \tag{A.6}$$

$$= \langle \mathrm{e}^{-\lambda\phi(\mathcal{C},\sigma_\tau)}|\mathcal{T}\mathrm{e}^{\int_0^\tau \mathbb{W}_\lambda}|\mathrm{e}^{-(1-\lambda)\phi(\mathcal{C},\sigma_0)}\rangle \tag{A.7}$$

where generically $|f(\mathcal{C})\rangle$ denotes the vector of components $f(\mathcal{C})$. The time-ordered exponential

$$\mathcal{T}\exp\left(\int_0^t \mathbb{W}_\lambda\right) = \sum_{n\ge 0}\int_0^t \mathrm{d}t_n \int_0^{t_n}\mathrm{d}t_{n-1}\cdots\int_0^{t_2}\mathrm{d}t_1\,\mathbb{W}_\lambda(t_n)\cdots\mathbb{W}_\lambda(t_1) \tag{A.8}$$

solves $\partial_t|\psi(t)\rangle = \mathbb{W}_\lambda(t)|\psi(t)\rangle$ as is easily checked.

A.2. A primer: the case $\lambda = 1$ (original Hatano–Sasa equality)

From (A.7) one has $\langle \mathrm{e}^{-\mathcal{Y}}\rangle = \langle \mathrm{e}^{-\phi(\mathcal{C},\sigma_\tau)}|\mathcal{T}\mathrm{e}^{\int_0^\tau \mathbb{W}_{\lambda=1}}|1\rangle$ but

$$(\mathbb{W}_{\lambda=1}|1\rangle)_{\mathcal{C}} = \sum_{\mathcal{C}'}(\mathbb{W}_{\lambda=1})_{\mathcal{C}\mathcal{C}'} = \sum_{\mathcal{C}'}[\mathrm{e}^{-[\phi(\mathcal{C}',\sigma)-\phi(\mathcal{C},\sigma)]}W(\mathcal{C}'\to\mathcal{C},\sigma) - \delta_{\mathcal{C},\mathcal{C}'}r(\mathcal{C},\sigma)] = 0 \tag{A.9}$$

since by definition of a steady state

$$\sum_{\mathcal{C}'}[\mathrm{e}^{-\phi(\mathcal{C}',\sigma)}W(\mathcal{C}'\to\mathcal{C},\sigma) - \delta_{\mathcal{C},\mathcal{C}'}r(\mathcal{C},\sigma)\mathrm{e}^{-\phi(\mathcal{C},\sigma)}] = 0 \tag{A.10}$$

which means that $\langle \mathrm{e}^{-\mathcal{Y}}\rangle = 1$ as expected. Note that the derivation we have presented makes no use of time discretization or chain rules compared to [9].

A.3. Symmetry for the Hatano–Sasa large deviation function

Starting from (A.7) one has (denoting A^{T} the matrix transpose of A)

$$\langle \mathrm{e}^{-\lambda\mathcal{Y}}\rangle = \langle \mathrm{e}^{-(1-\lambda)\phi(\mathcal{C},\sigma_0)}|(\mathcal{T}\mathrm{e}^{\int_0^\tau \mathbb{W}_\lambda})^{\mathrm{T}}|\mathrm{e}^{-\lambda\phi(\mathcal{C},\sigma_\tau)}\rangle. \tag{A.11}$$

One checks, for example, from (A.8) that the transpose of the time-ordered exponential is

$$(\mathcal{T}\mathrm{e}^{\int_0^\tau \mathbb{W}_\lambda})^{\mathrm{T}} = \mathcal{T}\mathrm{e}^{\int_0^\tau (\mathbb{W}_\lambda^{\mathrm{R}})^{\mathrm{T}}} \tag{A.12}$$

where $\mathbb{W}_\lambda^{\mathrm{R}}$ is the operator of evolution with the time-reversed protocol $\sigma^{\mathrm{R}}(t) \equiv \sigma(\tau - t)$. Besides

$$((\mathbb{W}_\lambda^{\mathrm{R}})^{\mathrm{T}})_{\mathcal{C}\mathcal{C}'} = (\mathbb{W}_\lambda^{\mathrm{R}})_{\mathcal{C}'\mathcal{C}} \tag{A.13}$$

$$= \mathrm{e}^{-\lambda[\phi(\mathcal{C},\sigma^{\mathrm{R}})-\phi(\mathcal{C}',\sigma^{\mathrm{R}})]} W(\mathcal{C} \to \mathcal{C}', \sigma^{\mathrm{R}}) - \delta_{\mathcal{C},\mathcal{C}'} r(\mathcal{C},\sigma) \tag{A.14}$$

$$= \mathrm{e}^{-(1-\lambda)[\phi(\mathcal{C}',\sigma^{\mathrm{R}})-\phi(\mathcal{C},\sigma^{\mathrm{R}})]} W^\dagger(\mathcal{C}' \to \mathcal{C}, \sigma^{\mathrm{R}}) - \delta_{\mathcal{C},\mathcal{C}'} r^\dagger(\mathcal{C},\sigma) \tag{A.15}$$

which shows that

$$(\mathbb{W}_\lambda^{\mathrm{R}})^{\mathrm{T}} = (\mathbb{W}_{1-\lambda}^{\mathrm{R}})^\dagger. \tag{A.16}$$

Using now (A.11) and (A.12), one arrives at

$$\langle \mathrm{e}^{-\lambda\mathcal{Y}} \rangle = \langle \mathrm{e}^{-(1-\lambda)\phi(\mathcal{C},\sigma_\tau^{\mathrm{R}})} | \mathcal{T} \mathrm{e}^{\int_0^\tau (\mathbb{W}_{1-\lambda}^{\mathrm{R}})^\dagger} | \mathrm{e}^{-\lambda\phi(\mathcal{C},\sigma_0^{\mathrm{R}})} \rangle. \tag{A.17}$$

Since the dynamics of rates W and $W^\dagger$ have the same steady state and hence the same ϕ, we see comparing with (A.7) that the large deviation function (ldf) at s for the protocol σ is the same as the ldf at $1-\lambda$ for the time-reversed protocol σ^{R} and the dual dynamics:

$$\mu(\lambda, \tau; \sigma) = \mu^\dagger(1-\lambda, \tau; \sigma^{\mathrm{R}}). \tag{A.18}$$

A.4. Connection with Lebowitz–Spohn-like current

In [6] Lebowitz and Spohn introduced a history-dependent observable Q obeying a Gallavotti–Cohen-like symmetry for Markov processes with *time-independent* jump rates. We may generalize their approach by defining, for a given history and a fixed time-dependent protocol σ:

$$Q_\tau = \sum_{k=1}^{K} \log \frac{W(\mathcal{C}_{k-1} \to \mathcal{C}_k, \sigma_{t_k})}{W^\dagger(\mathcal{C}_{k-1} \to \mathcal{C}_k, \sigma_{t_k})} = -\hat{h}_\tau + \sum_{k=1}^{K} \log \frac{W(\mathcal{C}_{k-1} \to \mathcal{C}_k, \sigma_{t_k})}{W(\mathcal{C}_k \to \mathcal{C}_{k-1}, \sigma_{t_k})} \tag{A.19}$$

where $\hat{h}_\tau$ is the reduced Hatano–Sasa defined in (113). Note that, when the protocol is time-independent, the contribution $\hat{h}_\tau$ to Q merely sums up to $\phi(\mathcal{C}_K) - \phi(\mathcal{C}_0)$ and one recovers the Lebowitz–Spohn definition of Q. Moreover, detailed balance is equivalent to having $Q_\tau = 0$ for all histories. One studies here the symmetries of $\langle \mathrm{e}^{-\lambda\mathcal{Y}-\kappa Q_\tau} \rangle$.

In the same way as above, one may introduce the probability $P(\mathcal{C}, \hat{h}, Q, t)$ of being in configuration $\mathcal{C}$ at time t, having observed a value $\hat{h}$ of the reduced Hatano–Sasa functional and a value Q of the Lebowitz–Spohn one—having started from configuration $\mathcal{C}_0$. Following the same steps, one checks that the Laplace transform

$$\hat{P}(\mathcal{C}, \lambda, \kappa, t) = \int \mathrm{d}\hat{h}\, \mathrm{d}Q\, \mathrm{e}^{-\lambda\hat{h}-\kappa Q} P(\mathcal{C}, \hat{h}, Q, t) \tag{A.20}$$

evolves according to $\partial_t |\hat{P}(\lambda, \kappa, t)\rangle = \mathbb{W}_{\lambda,\kappa} |\hat{P}(\lambda, \kappa, t)\rangle$ with a time-dependent operator $\mathbb{W}_{\lambda,\kappa}$ of elements

$$(\mathbb{W}_{\lambda,\kappa})_{\mathcal{C},\mathcal{C}'} = \mathrm{e}^{-(\lambda-\kappa)[\phi(\mathcal{C}',\sigma)-\phi(\mathcal{C},\sigma)]} W(\mathcal{C}' \to \mathcal{C}, \sigma)^{1-\kappa} W(\mathcal{C}' \to \mathcal{C}, \sigma)^{\kappa} - \delta_{\mathcal{C},\mathcal{C}'} r(\mathcal{C},\sigma). \tag{A.21}$$

J. Stat. Mech. (2012) P02009

One checks that upon transposition the operator possesses the following symmetry:

$$(\mathbb{W}_{\lambda,\kappa})^{\mathrm{T}} = \mathbb{W}_{1-\lambda,1-\kappa}. \tag{A.22}$$

Using now that as in (A.7) our quantity of interest $\langle \mathrm{e}^{-\lambda\mathcal{Y}-\kappa Q_\tau}\rangle$ is

$$\langle \mathrm{e}^{-\lambda\mathcal{Y}-\kappa Q_\tau}\rangle = \sum_{\mathcal{C}_0} P_{\mathrm{st}}(\mathcal{C}_0;\sigma_0)\int \mathrm{d}\hat{h}\,\mathrm{d}Q \sum_{\mathcal{C}} P(\mathcal{C},\hat{h},Q,\tau)\mathrm{e}^{-\lambda\hat{h}-\kappa Q}\mathrm{e}^{-\lambda[\phi(\mathcal{C}_0,\sigma_0)-\phi(\mathcal{C},\sigma_\tau)]} \tag{A.23}$$

$$= \langle \mathrm{e}^{-\lambda\phi(\mathcal{C},\sigma_\tau)}|\mathcal{T}\mathrm{e}^{\int_0^\tau \mathbb{W}_{\lambda,\kappa}}|\mathrm{e}^{-(1-\lambda)\phi(\mathcal{C},\sigma_0)}\rangle \tag{A.24}$$

and using (A.11) and (A.12) one obtains that the large deviation function $\mu(\lambda,\kappa,t)$ defined as $\mu(\lambda,\kappa,\tau;\sigma) = \log\langle \mathrm{e}^{-\lambda\mathcal{Y}-\kappa Q_\tau}\rangle$ possesses the symmetry (at all times)

$$\mu(\lambda,\kappa,\tau;\sigma) = \mu(1-\lambda,1-\kappa,\tau;\sigma^{\mathrm{R}}). \tag{A.25}$$

Note first that no reference is made here to the dual dynamics. This also gives

$$\langle \mathrm{e}^{-\lambda\mathcal{Y}-\kappa Q_\tau}\rangle = \langle \mathrm{e}^{-(1-\lambda)\mathcal{Y}-(1-\kappa)Q_\tau}\rangle_{\sigma^{\mathrm{R}}}. \tag{A.26}$$

Moreover, focusing on the case $\kappa = 0$ one sees that

$$\langle \mathrm{e}^{-\lambda\mathcal{Y}}\rangle = \langle \mathrm{e}^{-(1-\lambda)\mathcal{Y}-Q}\rangle_{\sigma^{\mathrm{R}}} \tag{A.27}$$

which is the equivalent of the equality

$$S_\lambda[\mathbf{x}^{\mathrm{R}};\sigma] = S_{1-\lambda}[\mathbf{x};\sigma^{\mathrm{R}}] + \beta Q_{\mathrm{hk}}[\mathbf{x};\sigma^{\mathrm{R}}] \tag{A.28}$$

valid for systems with Langevin dynamics. In the same way, for $\lambda = 1$ one sees that

$$\langle \mathrm{e}^{-Q}\rangle = 1. \tag{A.29}$$

This shows that Q plays the same role for Markov chains as βQ_{hk} for systems with Langevin dynamics.

References

[1] Bustamante C, Liphardt J and Ritort F, *The nonequilibrium thermodynamics of small systems*, 2005 *Phys. Today* **58** 43

[2] Ritort F, *Nonequilibrium fluctuations in small systems: from physics to biology*, 2008 *Adv. Chem. Phys.* **137** 31

[3] Evans D J, Cohen E G D and Morriss G P, *Probability of second law violations in shearing steady states*, 1993 *Phys. Rev. Lett.* **71** 2401

[4] Gallavotti G and Cohen E G D, *Dynamical ensembles in nonequilibrium statistical mechanics*, 1995 *Phys. Rev. Lett.* **74** 2694

[5] Kurchan J, *Fluctuation theorem for stochastic dynamics*, 1998 *J. Phys. A: Math. Gen.* **31** 3719

[6] Lebowitz J L and Spohn H, *A Gallavotti–Cohen type symmetry in the large deviation functional for stochastic dynamics*, 1999 *J. Stat. Phys.* **95** 333

[7] Jarzynski C, *Nonequilibrium equality for free energy differences*, 1997 *Phys. Rev. Lett.* **78** 2690
Jarzynski C, *Equilibrium free-energy differences from nonequilibrium measurements: a master-equation approach*, 1997 *Phys. Rev.* E **56** 5018

[8] Crooks G E, *Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems*, 1998 *J. Stat. Phys.* **90** 1481
Crooks G E, *Path-ensemble averages in systems driven far from equilibrium*, 2000 *Phys. Rev.* E **61** 2361

[9] Hatano T and Sasa S-I, *Steady-state thermodynamics of Langevin systems*, 2001 *Phys. Rev. Lett.* **86** 3463

[10] Speck T and Seifert U, *Integral fluctuation theorem for the housekeeping heat*, 2005 *J. Phys. A: Math. Gen.* **38** 581

J. Stat. Mech. (2012) P02009

[11] Seifert U, *Entropy production along a stochastic trajectory and an integral fluctuation theorem*, 2005 *Phys. Rev. Lett.* **95** 040602

[12] Horowitz J M and Vaikuntanathan S, *Nonequilibrium detailed fluctuation theorem for repeated discrete feedback*, 2010 *Phys. Rev.* E **82** 061120

Horowitz J M and Parrondo J M R, *Thermodynamic reversibility in feedback processes*, 2011 *Europhys. Lett.* **95** 10005

[13] Sagawa T and Ueda M, *Generalized Jarzynski equality under nonequilibrium feedback control*, 2010 *Phys. Rev. Lett.* **104** 090602

[14] Abreu D and Seifert U, *Extracting work from a single heat bath through feedback*, 2011 *Europhys. Lett.* **94** 10001

[15] Kurchan J, *Non-equilibrium work relations*, 2007 *J. Stat. Mech.* P07005

[16] Harris R J and Schütz G M, *Fluctuation theorems for stochastic dynamics*, 2007 *J. Stat. Mech.* P07020

[17] Zamponi F, Bonetto F, Cugliandolo L F and Kurchan J, *A fluctuation theorem for non-equilibrium relaxational systems driven by external forces*, 2005 *J. Stat. Mech.* P09013

[18] Aron C, Biroli G and Cugliandolo L F, *Symmetries of generating functionals of Langevin processes with colored multiplicative noise*, 2010 *J. Stat. Mech.* P11018

[19] Esposito M and Van den Broeck C, *Three detailed fluctuation theorems*, 2010 *Phys. Rev. Lett.* **104** 090601

[20] Chernyak V Y, Chertkov M and Jarzynski C, *Path-integral analysis of fluctuation theorems for general Langevin processes*, 2006 *J. Stat. Mech.* P08001

[21] Esposito M and Van den Broeck C, *Three faces of the second law: I. Master equation formulation*, 2010 *Phys. Rev.* E **82** 011143

[22] Esposito M and Van den Broeck C, *The three faces of the second law: II. Fokker–Planck formulation*, 2010 *Phys. Rev.* E **82** 011144

[23] Oono Y and Paniconi M, *Steady state thermodynamics*, 1998 *Prog. Theor. Phys. Suppl.* **130** 29

[24] Hurtado P I, Perez-Espigares C, del Pozo J J and Garrido P L, *Symmetries in fluctuations far from equilibrium*, 2011 *Proc. Nat. Acad. Sci.* **108** 7704

[25] García-García R, Domínguez D, Lecomte V and Kolton A B, *Unifying approach for fluctuation theorems from joint probability distributions*, 2010 *Phys. Rev.* E **82** 030104(R)

[26] Seifert U, *Stochastic thermodynamics: principles and perspectives*, 2008 *Eur. J. Phys.* B **64** 423

[27] Gomez-Marin A, Parrondo J M R and Van den Broeck C, *The 'footprints" of irreversibility*, 2008 *Europhys. Lett.* **82** 50002

[28] Pérez-Espigares C, Kolton A B and Kurchan J, *An infinite family of Second Law-like inequalities*, 2011 arXiv:1110.0967v2

[29] Andrieux D and Gaspard P, *Fluctuation theorem for currents and Schnakenberg network theory*, 2007 *J. Stat. Phys.* **127** 107

[30] Lau A W C and Lubensky T C, *State-dependent diffusion: thermodynamic consistency and its path integral formulation*, 2007 *Phys. Rev.* E **76** 011123

[31] Trepagnier E H, Jarzynski C, Ritort F, Crooks G E, Bustamante C J and Liphardt J, *Experimental test of Hatano and Sasa's nonequilibrium steady-state equality*, 2004 *Proc. Nat. Acad. Sci.* **101** 15038

[32] Richardson J, *Noise in driven systems*, 1955 *IRE Trans. Inf. Theory* **1** 62

[33] Agarwal G S, *Fluctuation-dissipation theorems for systems in non-thermal equilibrium and applications*, 1972 *Z. Phys.* **252** 25

[34] Risken H, 1984 *The Fokker–Planck Equation* (Berlin: Springer) chapter 7

[35] Falcioni M, Isola S and Vulpiani A, *Correlation functions and relaxation properties in chaotic dynamics and statistical mechanics*, 1990 *Phys. Lett.* A **144** 341

[36] Prost J, Joanny J-F and Parrondo J M R, *Generalized fluctuation-dissipation theorem for steady-state systems*, 2009 *Phys. Rev. Lett.* **103** 090601

[37] Baiesi M, Maes C and Wynants B, *Fluctuations and response of nonequilibrium states*, 2009 *Phys. Rev. Lett.* **103** 010602

[38] Verley G, Mallick K and Lacoste D, *Modified fluctuation-dissipation theorem for non-equilibrium steady states and applications to molecular motors*, 2011 *Europhys. Lett.* **93** 10002

[39] Verley G, Chétrite R and Lacoste D, *Modified fluctuation-dissipation theorem near non-equilibrium states and applications to the Glauber–Ising chain*, 2011 *J. Stat. Mech.* P10025

[40] Marini B, Marconi U, Puglisi A, Rondoni L and Vulpiani A, *Fluctuation-dissipation: response theory in statistical physics*, 2008 *Phys. Rep.* **461** 111

[41] Lecomte V, Appert-Rolland C and van Wijland F, *Thermodynamic formalism for systems with Markov dynamics*, 2007 *J. Stat. Phys.* **127** 51

[42] Garrahan J P, Jack R L, Lecomte V, Pitard E, van Duijvendijk K and van Wijland F, *Dynamic first-order phase transition in kinetically constrained models of glasses*, 2007 *Phys. Rev. Lett.* **98** 195702

[43] Garrahan J P, Jack R L, Lecomte V, Pitard E, van Duijvendijk K and van Wijland F, *First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories*, 2009 *J. Phys. A: Math. Theor.* **42** 075007

[44] Turci F and Pitard E, *Large deviations and heterogeneities in a driven kinetically constrained model*, 2011 *Europhys. Lett.* **94** 10003

[45] Gorissen M and Vanderzande C, *Finite size scaling of current fluctuations in the totally asymmetric exclusion process*, 2011 *J. Phys. A: Math. Theor.* **44** 115005

[46] Bodineau T and Derrida B, *Cumulants and large deviations of the current through non-equilibrium steady states*, 2007 *C. R. Phys.* **8** 540

[47] Nemoto T and Sasa S-I, *Thermodynamic formula for the cumulant generating function of time-averaged current*, 2011 *Phys. Rev.* E **84** 061113

[48] Giardinà C, Kurchan J and Peliti L, *Direct evaluation of large-deviation functions*, 2006 *Phys. Rev. Lett.* **96** 120603

[49] Lecomte V and Tailleur J, *A numerical approach to large deviations in continuous time*, 2007 *J. Stat. Mech.* P03004

[50] Giardinà C, Kurchan J, Lecomte V and Tailleur J, *Simulating rare events in dynamical processes*, 2011 *J. Stat. Phys.* **145** 787

[51] Autieri E, Faccioli P, Sega M, Pederiva F and Orland H, *Dominant reaction pathways in high dimensional systems*, 2009 *J. Chem. Phys.* **130** 064106

[52] Bodineau T and Lefevere R, *Large deviations of lattice Hamiltonian dynamics coupled to stochastic thermostats*, 2008 *J. Stat. Phys.* **133** 1

[53] Maes C, Netočný K and Wynants B, *On and beyond entropy production the case of Markov jump processes*, 2008 *Markov Process. Relat. Fields* **14** 445

[54] Pitard E, Lecomte V and van Wijland F, *Dynamic transition in an atomic glass former: a molecular dynamics evidence*, 2011 *Europhys. Lett.* **96** 56002

[55] van Zon R *et al*, *Power and heat fluctuation theorems for electric circuits*, 2004 *Phys. Rev. Lett.* **92** 130601
Schuler S *et al*, *Experimental test of the fluctuation theorem for a driven two-level system with time-dependent rates*, 2005 *Phys. Rev. Lett.* **94** 180602
Tietz C *et al*, *Measurement of stochastic entropy production*, 2006 *Phys. Rev. Lett.* **97** 050602
Andrieux D *et al*, *Entropy production and time asymmetry in nonequilibrium fluctuations*, 2007 *Phys. Rev. Lett.* **98** 150601
Liphardt J *et al*, *Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality*, 2002 *Science* **296** 1832
Collin D *et al*, *Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies*, 2005 *Nature* **437** 231

[56] García-García R, 2012 unpublished

[57] Golubev D S, Utsumi Y, Marthaler M and Schön M, *Fluctuation theorem for a double quantum dot coupled to a point-contact electrometer*, 2011 *Phys. Rev.* B **84** 075323
Cuetara G B, Esposito M and Gaspard P, *Fluctuation theorems for capacitively coupled electronic currents*, 2011 *Phys. Rev.* B **84** 165114

[58] Puglisi A and Villamaina D, *Irreversible effects of memory*, 2009 *Europhys. Lett.* **88** 30004

J. Stat. Mech. (2012) P02009