RAPID COMMUNICATIONS

PHYSICAL REVIEW E **82**, 030104(R) (2010)

Unifying approach for fluctuation theorems from joint probability distributions

Reinaldo García-García,[1] Daniel Domínguez,[1] Vivien Lecomte,[2,3] and Alejandro B. Kolton[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, 2 Place Jussieu, 75005 Paris, France*
[3]*DPMC-MaNEP, Université de Genève, 24 Quai Ernest-Ansermet, 1211 Geneva 4, Switzerland*
(Received 12 July 2010; published 22 September 2010)

Any decomposition of the total trajectory entropy production for Markovian systems has a joint probability distribution satisfying a generalized detailed fluctuation theorem, when all the contributing terms are odd with respect to time reversal. The expression of the result does not bring into play dual probability distributions, hence easing potential applications. We show that several fluctuation theorems for perturbed nonequilibrium steady states are unified and arise as particular cases of this general result. In particular, we show that the joint probability distribution of the system and reservoir trajectory entropies satisfy a detailed fluctuation theorem valid for all times.

DOI: 10.1103/PhysRevE.82.030104 PACS number(s): 05.70.Ln, 05.40.−a

The nonequilibrium stochastic thermodynamics of small systems has attracted a lot of attention. From the experimental side the development of techniques for microscopic manipulation has allowed one to study fluctuations in small systems [1,2]. From the theoretical side a group of exact relations known as fluctuation theorems (FTs) [3–8] has shed light into the principles governing dissipation and fluctuations in nonequilibrium phenomena, as in driven systems in contact with thermal bath. Formally, the generality of the FTs can be attributed to the way probability distribution functions of different observables behave under time-reversal symmetry-breaking perturbations (see [9,10] for reviews on FT).

Oono and Paniconi [11] proposed a phenomenological framework for a "nonequilibrium steady-state (NESS) thermodynamics" aimed at describing fluctuating systems subjected to an external protocol. In this approach, the total exchanged heat during a time interval τ by a system initially prepared in a NESS is expressed as the sum of two contributions: $Q^{\mathrm{tot}}=Q^{\mathrm{ex}}+Q^{\mathrm{hk}}$. The "excess heat" Q^{ex} is, in average, associated with the energy exchange during transitions between steady states while the "housekeeping heat" Q^{hk} corresponds, in average, to the energy we need to supply to maintain the system in a NESS. Hatano and Sasa [12] introduced a formal framework for these phenomenological ideas and derived a FT which extends the second law of thermodynamics for transitions between NESSs controlled by external parameters $\sigma(t)$. The Hatano and Sasa FT is applicable to trajectories $x(t)$ evolved with a Langevin or more generally a Markovian dynamics starting from an initial condition sampled from a NESS compatible with the initial values $\sigma(0)$ of the control parameters.

Under identical conditions different FTs were subsequently proposed for NESS, involving the above decomposition of the exchanged heat. We can distinguish between the so-called integral and detailed FTs for systems initially prepared as described above. The integral fluctuation theorems (IFTs) are exact relations for the average over histories of different stochastic trajectory functionals $W[x]$, such as $\langle e^{-W}\rangle=1$. Examples are the Jarzynski FT for the total work [7], the Hatano-Sasa FT [12], and the Speck-Seifert FT [13]. The so-called detailed FTs (DFTs) are, on the other hand, exact relations for the probability distribution functions (PDFs) of different observables W, such as $P(W)/P^{\mathrm{R}}(-W)=e^{W}$, where $P^{\mathrm{R}}(W)$ corresponds to the time-reversed protocol $\sigma^{\mathrm{R}}(t)=\sigma(\tau-t)$, and a NESS initial condition compatible with $\sigma(\tau)$. Examples are given by the Crooks relation [8] or Seifert fluctuation theorem [14]. While observables satisfying a DFT trivially satisfy an IFT, the opposite is not always true. In many recently formulated DFTs, a modified "dual" PDF $P^{\dagger R}(W)$ enters into play [15,16], which corresponds to trajectories in a system with same stationary PDF but with reversed steady probability current [i.e., such that $P^{\dagger R}(W)=P^{R}(W)$ when detailed balance holds]. In general the dual dynamics is different from the original dynamics of the system. Thus, the presence of dual distributions is a strong limitation to the experimental use of such a DFT. A central result of our work is that generalized DFTs can be established *without relying on dual probabilities* by considering *joint* probability distributions for different complementary observables, instead of single PDFs. The joint probability distributions arise naturally from the above mentioned separation of two contributions to the total heat exchanged in a NESS. From this joint DFT all the known DFTs and IFTs follow in a straightforward way.

Among the fluctuation theorems formulated for Markovian dynamics, the total trajectory entropy production $\mathcal{S}[x;\sigma]=\ln(\mathcal{P}[x;\sigma]/\mathcal{P}[x^{\mathrm{R}};\sigma^{\mathrm{R}}])$ plays a fundamental role [5,6,14]. Here, $\mathcal{P}[x;\sigma]$ ($\mathcal{P}[x^{\mathrm{R}};\sigma^{\mathrm{R}}]$) denotes the probability density of trajectory x (time-reversed trajectory x^{R}) in the forward (backward) protocol. We include in $\mathcal{P}$ the initial distribution of x. We omit hereafter the final time τ in all trajectory functionals and use calligraphic symbols to denote functionals and normal symbols to denote their values. The total trajectory entropy production is odd upon time reversal: $\mathcal{S}[x^{\mathrm{R}};\sigma^{\mathrm{R}}]=-\mathcal{S}[x;\sigma]$. We show that any decomposition of $\mathcal{S}$ in M distinct contributions, $\mathcal{S}[x;\sigma]=\Sigma_i^M\mathcal{A}_i[x;\sigma]$, each of them being odd $\mathcal{A}_i[x^{\mathrm{R}};\sigma^{\mathrm{R}}]=-\mathcal{A}_i[x;\sigma]$, has a generating function satisfying the symmetry

$$\left\langle \exp\left(-\sum_i^M \lambda_i \mathcal{A}_i[x;\sigma]\right)\right\rangle = \left\langle \exp\left(-\sum_i^M (1-\lambda_i)\mathcal{A}_i[x;\sigma^{\mathrm{R}}]\right)\right\rangle_{\mathrm{R}}, \quad (1)$$

where λ_i are arbitrary parameters and $\langle \ldots \rangle_{\mathrm{R}}$ denotes average over trajectories in the reversed protocol [17]. This symmetry is equivalent to the following generalized DFT for the joint probability of $\mathcal{A}_i[x;\sigma]$'s:

$$\frac{P(A_1,A_2,\ldots,A_M)}{P^{\mathrm{R}}(-A_1,-A_2,\ldots,-A_M)} = e^S \quad \text{with } S=\sum_{i=1}^M A_i. \quad (2)$$

Note that the result involves no use of dual PDFs. To prove Eq. (1) we start by noting that the average of any observable $\mathcal{O}[x]$ over trajectories satisfies the relation

$$\begin{aligned}\langle \mathcal{O}[x;\sigma]\rangle &= \int \mathcal{D}x\mathcal{P}[x^{\mathrm{R}};\sigma^{\mathrm{R}}]\mathcal{O}[x;\sigma]e^{-\mathcal{S}[x^{\mathrm{R}};\sigma^{\mathrm{R}}]} \\ &= \int \mathcal{D}x\mathcal{P}[x;\sigma^{\mathrm{R}}]\mathcal{O}[x^{\mathrm{R}};\sigma]e^{-\mathcal{S}[x;\sigma^{\mathrm{R}}]} \\ &= \langle \mathcal{O}[x^{\mathrm{R}};\sigma]e^{-\mathcal{S}[x;\sigma^{\mathrm{R}}]}\rangle_{\mathrm{R}}, \end{aligned} \quad (3)$$

where we have used $\mathcal{S}[x;\sigma]=-\mathcal{S}[x^{\mathrm{R}};\sigma^{\mathrm{R}}]$ together with the change of variable $x \to x^{\mathrm{R}}$. Considering $\mathcal{S}[x;\sigma]=\Sigma_i^M \mathcal{A}_i[x;\sigma]$ the proof comes around:

$$\begin{aligned}\langle \exp(-\Sigma_i^M \lambda_i \mathcal{A}_i[x;\sigma])\rangle &= \langle \exp(-\Sigma_i^M \lambda_i \mathcal{A}_i[x^{\mathrm{R}};\sigma] - \mathcal{S}[x;\sigma^{\mathrm{R}}])\rangle_{\mathrm{R}} \\ &= \langle \exp\{-\Sigma_i^N (1-\lambda_i)\mathcal{A}_i[x;\sigma^{\mathrm{R}}]\}\rangle_{\mathrm{R}}.\end{aligned}$$

Equation (2) is proved in a similar way, or also follows from Eq. (1) since it is a symmetry for the generating function of the joint distribution $P(A_1,\ldots,A_M)$. Before considering their particular application for explicit decompositions of $\mathcal{S}$ we note that symmetries (1) and (2) are valid for all times τ for systems prepared in any initial distribution $P_i(x(0))$. In particular, we see that the *total trajectory entropy production FTs* $\langle e^{-\mathcal{S}}\rangle=1$ and $P(S)/P^{\mathrm{R}}(-S)=e^S$ hold without further assumption.

We will consider two generic frameworks where our result applies: systems described (i) by Langevin dynamics and (ii) by a Markovian dynamics on discrete configurations, exemplifying their parallel features. First we consider a particle driven by a constant force f in a potential U,

$$\dot{x} = -\partial_x U(x;\alpha(t)) + f + \xi, \quad (4)$$

where $\alpha(t)$ represents a set of control parameters and $\xi(t)$ is the Gaussian white noise, with $\langle \xi(t)\rangle=0$ and $\langle \xi(t)\xi(t')\rangle=2T\delta(t-t')$ due to a thermal bath at temperature T. For example, this system can be out of equilibrium when U has periodic boundary conditions. We consider for simplicity a single degree of freedom x, but our results are easily generalized, e.g., in larger dimensions and/or with more particles. For a stochastic trajectory generated by Eq. (4) we define the total exchanged heat as [18]

$$\beta\mathcal{Q}^{\mathrm{tot}}[x;\sigma] = -\beta\int_0^\tau dt\dot{x}[\partial_x U(x;\alpha) - f]. \quad (5)$$

The total exchanged heat in a trajectory can be split as $\mathcal{Q}^{\mathrm{tot}}=\mathcal{Q}^{\mathrm{hk}}+\mathcal{Q}^{\mathrm{ex}}$ [11]. Defining $\phi(x;\sigma)=-\ln \rho_{\mathrm{SS}}(x,\sigma)$ from the steady-state probability density $\rho_{\mathrm{SS}}(x,\sigma)$ at fixed values of $\sigma=(\alpha,f)$ Hatano and Sasa [12] proposed

$$\beta\mathcal{Q}^{\mathrm{hk}}[x;\sigma] = \int_0^\tau dt\dot{x}\{\partial_x\phi(x;\sigma) - \beta[\partial_x U(x;\alpha) - f]\}, \quad (6)$$

for the housekeeping heat, and

$$\beta\mathcal{Q}^{\mathrm{ex}}[x;\sigma] = -\int_0^\tau dt\dot{x}\partial_x\phi(x;\sigma), \quad (7)$$

for the excess heat. The Hatano-Sasa functional [12] $\mathcal{Y}[x;\sigma]$ is then defined as

$$\mathcal{Y}[x;\sigma] \equiv \int_0^\tau dt\dot{\sigma}\partial_\sigma\phi(x;\sigma) = \beta\mathcal{Q}^{\mathrm{ex}}[x;\sigma] + \Delta\phi(x;\sigma), \quad (8)$$

where $\Delta\phi(x;\sigma)=\phi(x(\tau);\sigma(\tau))-\phi(x(0);\sigma(0))$ is a time boundary term.

We now assume that the system is initially prepared with a distribution corresponding to a NESS compatible with $\sigma(0)$, $P_i(x(0))=\rho_{\mathrm{SS}}(x(0),\sigma(0))$. With the previous definitions of Eqs. (5)–(8) it is known that the total trajectory entropy production $\mathcal{S}$ can be decomposed as the sum of two contributions, in two different ways [10,11,13,16]:

$$\mathcal{S} = \mathcal{Y} + \beta\mathcal{Q}^{\mathrm{hk}} = \Delta\phi + \beta\mathcal{Q}^{\mathrm{tot}}. \quad (9)$$

Similar decompositions also exist for Markovian dynamics: we consider now discrete configurations $\{\mathcal{C}\}$ with transition rates $W(\mathcal{C}\to\mathcal{C}';\sigma)$ between configurations. They depend on σ, an external control parameter which may vary in time. The probability density at time t obeys the Markovian dynamics $\partial_t P(\mathcal{C},t)=\Sigma_{\mathcal{C}'}W(\mathcal{C}'\to\mathcal{C};\sigma(t))P(\mathcal{C}',t) - r(\mathcal{C};\sigma(t))P(\mathcal{C},t)$, where $r(\mathcal{C};\sigma)=\Sigma_{\mathcal{C}'}W(\mathcal{C}\to\mathcal{C}';\sigma)$ is the escape rate from configuration $\mathcal{C}$. A history of the system is described by the succession of configuration $(\mathcal{C}_0,\ldots,\mathcal{C}_K)$ visited by the system, with $\mathcal{C}_k$ being visited between t_k and t_{k+1}. Starting from initial distribution $P_{\mathrm{i}}(\mathcal{C}(0),\sigma(0))$, the probability of a history is $\mathcal{P}[\mathcal{C};\sigma] =\exp[-\int_0^\tau dtr(\mathcal{C}(t);\sigma(t))]\Pi_{k=1}^K W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})P_{\mathrm{i}}(\mathcal{C}(0),\sigma(0))$, meaning that the mean value of an history-dependent observable $\mathcal{O}$ is given by $\langle\mathcal{O}\rangle =\Sigma_{K\ge 0}\Sigma_{\mathcal{C}_0\cdots\mathcal{C}_K}\int_0^t dt_K\cdots\int_0^{t_2}dt_1\mathcal{O}[\mathcal{C},\sigma]\mathcal{P}[\mathcal{C},\sigma]$. We obtain that the total trajectory entropy production $\mathcal{S}[\mathcal{C};\sigma] =\ln(\mathcal{P}[\mathcal{C};\sigma]/\mathcal{P}[\mathcal{C}^{\mathrm{R}};\sigma^{\mathrm{R}}])$ has a first decomposition $\mathcal{S}=\Delta\phi +\beta\mathcal{Q}^{\mathrm{tot}}$ with $\Delta\phi=\log[P_{\mathrm{i}}(\mathcal{C}(0),\sigma(0))/P_{\mathrm{i}}(\mathcal{C}(\tau),\sigma(\tau))]$ and

$$\beta\mathcal{Q}^{\mathrm{tot}} = \sum_{k=1}^K \log\frac{W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})}{W(\mathcal{C}_k\to\mathcal{C}_{k-1},\sigma_{t_k})}. \quad (10)$$

Although there is no natural definition of β we write $\beta\mathcal{Q}^{\mathrm{tot}}$ to exemplify the parallel with Langevin dynamics.

Turning to the first decomposition, let us now assume that the initial distribution is steady state: $P_{\mathrm{i}}=\rho_{\mathrm{SS}}=e^{-\phi}$. One directly checks that the Hatano-Sasa functional $\mathcal{Y}[\mathcal{C},\sigma]=\int_0^\tau dt\dot{\sigma}\partial_\sigma\phi$ can be written as

$$\mathcal{Y} = [\phi(\mathcal{C},\sigma)]_0^\tau - \sum_{k=1}^{K} [\phi(\mathcal{C}_k,\sigma_{t_k}) - \phi(\mathcal{C}_{k-1},\sigma_{t_k})]. \quad (11)$$

Besides, defining the housekeeping work as

$$\beta \mathcal{Q}^{\text{hk}}[\mathcal{C},\sigma] = \sum_{k=1}^{K} \log \frac{W(\mathcal{C}_{k-1}\to\mathcal{C}_k,\sigma_{t_k})}{W(\mathcal{C}_k\to\mathcal{C}_{k-1},\sigma_{t_k})} + \sum_{k=1}^{K} \phi(\mathcal{C}_k,\sigma_{t_k}) - \phi(\mathcal{C}_{k-1},\sigma_{t_k}), \quad (12)$$

we check that the decomposition $\mathcal{S}=\mathcal{Y}+\beta\mathcal{Q}^{\text{hk}}$ holds [19]. The parallel between Markovian and Langevin frameworks also appears by specializing to rates $W(k\to k\pm 1) = e^{-(\beta/2)(V_{k\pm1}-V_k)}$ of jumping on a one-dimensional lattice from site k to $k\pm 1$ in a tilted potential $V_k=U_k-kf$: in the continuum limit, one recovers the Langevin observables [19].

In the first decomposition, $\mathcal{Y}$ can be identified with the so-called nonadiabatic contribution (since it vanishes for quasistatic protocols) to the trajectory entropy $\mathcal{S}_{\text{na}}\equiv\mathcal{Y}$, while $\beta\mathcal{Q}^{\text{hk}}$ (which is continuously produced in the steady state) can be identified with the so-called adiabatic part $\mathcal{S}_{\text{a}} \equiv \beta\mathcal{Q}^{\text{hk}}$ [16]. On the other hand, in the second decomposition of $\mathcal{S}$, $\Delta\phi$ can be identified with the system contribution $\mathcal{S}_s\equiv\Delta\phi$, while $\beta\mathcal{Q}^{\text{tot}}$ can be identified with the reservoir contribution $\mathcal{S}_r\equiv\beta\mathcal{Q}^{\text{tot}}$ to the total trajectory entropy production.

The entropy decompositions of Eq. (9) satisfy the conditions for the application of identity (1) or (2) since each term is odd with respect to time reversal in both decompositions. We can thus write DFTs (valid for all times τ) for the joint probabilities as

$$\frac{P(Y,\beta Q^{\text{hk}})}{P^{\text{R}}(-Y,-\beta Q^{\text{hk}})} = e^{Y+\beta Q^{\text{hk}}}, \quad (13)$$

$$\frac{P(\Delta\phi,\beta Q^{\text{tot}})}{P^{\text{R}}(-\Delta\phi,-\beta Q^{\text{tot}})} = e^{\Delta\phi+\beta Q^{\text{tot}}}. \quad (14)$$

The corresponding IFTs can be written as

$$\langle e^{-\lambda\mathcal{Y}-\kappa\beta\mathcal{Q}^{\text{hk}}}\rangle = \langle e^{-(1-\lambda)\mathcal{Y}_{\text{R}}-(1-\kappa)\beta\mathcal{Q}_{\text{R}}^{\text{hk}}}\rangle_{\text{R}}, \quad (15)$$

$$\langle e^{-\lambda\Delta\phi-\kappa\beta\mathcal{Q}^{\text{tot}}}\rangle = \langle e^{-(1-\lambda)\Delta\phi_{\text{R}}-(1-\kappa)\beta\mathcal{Q}_{\text{R}}^{\text{tot}}}\rangle_{\text{R}}. \quad (16)$$

Here, $\mathcal{X}_{\text{R}}$ denotes $\mathcal{X}[x;\sigma^{\text{R}}]$, with $\mathcal{X}$ representing $\mathcal{Y}$, $\mathcal{Q}^{\text{hk}}$, $\mathcal{Q}^{\text{tot}}$, or $\Delta\phi$. From Eqs. (13) and (14) we have, in terms of S_s, S_r, S_a, and S_{na}, that $P(S_s,S_r)/P^{\text{R}}(-S_s,-S_r)=e^{S_s+S_r}$ and $P(S_a,S_{na})/P^{\text{R}}(-S_a,-S_{na})=e^{S_a+S_{na}}$. It is worth noting that these relations do not involve dual PDFs, and thus they can be tested for a physical system with a given dynamics. We also note that while one can show that S_a and S_{na} satisfy each one separately a DFT by using dual PDFs [16], S_s and S_r satisfy a joint DFT although they do not satisfy separately a DFT. (In the asymptotic long-time limit a DFT for S_r can be formulated for systems with bounded energy under steady-state conditions [3–6,14], while for systems with unbounded energies an "extended fluctuation relation" is necessary [20].)

Let us now derive from a unified view the known FTs. As an intermediate step, it is useful to define a dual trajectory weight $\mathcal{P}^\dagger[x]$ as [15,16]

$$\mathcal{P}^\dagger[x;\sigma] = \mathcal{P}[x;\sigma]e^{-\beta\mathcal{Q}^{\text{hk}}[x;\sigma]}. \quad (17)$$

For Markovian dynamics the dual probability $\mathcal{P}^\dagger$ corresponds to the dynamics in the so-called dual rates $W^\dagger(\mathcal{C}\to\mathcal{C}',\sigma) \equiv e^{-[\phi(\mathcal{C}',\sigma)-\phi(\mathcal{C},\sigma)]}W(\mathcal{C}'\to\mathcal{C},\sigma)$ which share the same steady state as the original dynamics. In the case of the Langevin dynamics of Eq. (4), it corresponds to trajectories generated by the equation $\dot{x}=-\partial_x U^\dagger(x;\alpha(t))+f^\dagger+\xi$, with $U^\dagger=2\phi/\beta-U$ and $f^\dagger=-f$. This equation also has the same steady state as the original one. Let us stress that the dual dynamics corresponds to trajectories in a different physical system. We now follow Eq. (17) to write the dual joint PDF related to Eq. (13) as

$$P^\dagger(Y,\beta Q^{\text{hk}}) = P(Y,-\beta Q^{\text{hk}})e^{\beta Q^{\text{hk}}}, \quad (18)$$

which is normalized. Integrating this relation over Y, we first obtain the DFT [16] $P(\beta Q^{\text{hk}})=e^{\beta Q^{\text{hk}}}P^\dagger(-\beta Q^{\text{hk}})$, and hence the IFT $\langle e^{-\beta Q^{\text{hk}}}\rangle=1$ [13]. Using now successively Eqs. (13) and (18),

$$P(Y) = e^Y\int d(\beta Q^{\text{hk}})e^{\beta Q^{\text{hk}}}P^{\text{R}}(-Y,-\beta Q^{\text{hk}}) = e^Y\int d(\beta Q^{\text{hk}})P^{\dagger R}(-Y,\beta Q^{\text{hk}}), \quad (19)$$

we see that the DFT $P(Y)=e^Y P^{\dagger R}(-Y)$ [15] holds. This implies the corresponding IFT $\langle e^{-\mathcal{Y}}\rangle=1$ [12] [also derived from setting $\lambda=1$ and $\kappa=0$ in Eq. (15) and using the Speck-Seifert IFT]. Finally, we note that although we have so far assumed steady-state initial conditions, relations (14) and (16) remain valid for an arbitrary initial distribution $P_{\text{i}}(\mathcal{C}(0),\sigma(0))$ by replacing $\Delta\phi\to\log[P_{\text{i}}(\mathcal{C}(\tau),\sigma(\tau))/P_{\text{i}}(\mathcal{C}(0),\sigma(0))]$. By taking $\lambda=\kappa=1$ in Eq. (16) we thus reobtain the IFT noted by Seifert [14].

As an illustration of our approach, let us show how joint FTs provide insights on the experimental error in the evaluation of entropy productions. Consider an experiment where the steady state can be evaluated for different values of the control parameter σ, e.g., microspheres optically driven in a liquid [21]. Having in hand an experimental evaluation ϕ_{exp} of ϕ, we write

$$\mathcal{S} = \mathcal{Y}_{\text{e}} + \delta\mathcal{Y} + \beta\mathcal{Q}^{\text{hk}}, \quad (20)$$

where $\mathcal{Y}_{\text{e}}[x;\sigma]=\int_0^\tau dt\dot{\sigma}\partial_\sigma\phi_{\text{exp}}$ and $\delta\mathcal{Y}=\mathcal{Y}-\mathcal{Y}_{\text{e}}$ is the difference between exact and experimental Hatano-Sasa functionals. Starting from the NESS associated with $\sigma(0)$ and assuming that each of the terms in Eq. (20) is odd upon time reversal, we can use Eq. (2) for $M=3$, obtaining

$$P(Y_{\text{e}},\delta Y,\beta Q^{\text{hk}}) = P^{\text{R}}(-Y_{\text{e}},-\delta Y,-\beta Q^{\text{hk}})e^{Y_{\text{e}}+\delta Y+\beta Q^{\text{hk}}}.$$

Using Eq. (18), this gives $P(Y_{\text{e}},\delta Y)=P^{\dagger\text{R}}(-Y_{\text{e}},-\delta Y)e^{Y_{\text{e}}+\delta Y}$, and hence also the IFT

$$\langle e^{-\mathcal{Y}_e}\rangle = \langle e^{-\delta\mathcal{Y}_R}\rangle_R^\dagger. \tag{21}$$

From the analysis of $\langle e^{-\delta\mathcal{Y}_R}\rangle_R^\dagger$ for specific cases of experimental errors, one can estimate the dispersion of the experimentally obtained $\langle e^{-\mathcal{Y}_e}\rangle$ around $\langle e^{-\mathcal{Y}}\rangle=1$ [19].

As a second example let us consider a system initially prepared in a NESS with a 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 case a modified fluctuation-dissipation theorem has been derived in [22], which relates dissipation under small perturbations around a NESS with fluctuations in the corresponding steady state. Expanding the exponentials in Eq. (15) up to second order in $\delta\sigma$ and then to second order in λ and order zero in κ, we arrive at (see [19] for details)

$$\langle \mathcal{B}_{ij}(0,\tau)\rangle_0 = \left\langle \mathcal{B}_{ij}(\tau,0)\exp\left[-\beta\int_0^\tau dt\dot{x}(t)v_s(x(t);\sigma^0)\right]\right\rangle_0, \tag{22}$$

where $\mathcal{B}_{ij}(t,t')=[\partial\phi(x(t);\sigma^0)/\partial\sigma_i][\partial\phi(x(t');\sigma^0)/\partial\sigma_j]$ and $v_s=\beta^{-1}\partial_x\phi-(\partial_x U-f)$ corresponds to the average velocity in the NESS associated with σ. For a Boltzmann-Gibbs steady state, this result reduces to the symmetry $\mathcal{C}_{ij}(\tau)=\mathcal{C}_{ij}(-\tau)$, with $\mathcal{C}_{ij}(\tau)=\langle\mathcal{B}_{ij}(0,\tau)\rangle_0$. Equation (22) can also be derived from Eq. (19) in Ref. [15]. With the use of the joint PDF we can obtain further new results. Let us introduce a weighted correlation function as

$$\mathcal{C}_{ij}^W(\tau,0) = \frac{\left\langle \mathcal{B}_{ij}(0,\tau)\exp\left[-\frac{\beta}{2}\int_0^\tau dt\dot{x}(t)v_s(x(t);\sigma^0)\right]\right\rangle_0}{\left\langle \exp\left[-\frac{\beta}{2}\int_0^\tau dt\dot{x}(t)v_s(x(t);\sigma^0)\right]\right\rangle_0}. \tag{23}$$

This correlation function carries explicit information about the lack of detailed balance and reduces to the usual one when the system is able to equilibrate. Using Eq. (15) with $\kappa=\frac{1}{2}$ and repeating the same procedure used to obtain Eq. (22) we arrive at the symmetry $\mathcal{C}_{ij}^W(\tau,0)=\mathcal{C}_{ij}^W(0,\tau)$, which reduces to the known result for equilibrium dynamics when detailed balance holds.

In conclusion, identities (1) and (2) and their immediate consequences for Markovian systems are the main results of this work. Equation (1) or Eq. (2) indeed contains, as particular cases, several known FTs such as the ones previously derived by Hatano and Sasa [12], Speck and Seifert [13], Seifert [14], Chernyak *et al.* [15], and Esposito and Van den Broeck [16]. In addition, an exact DFT, valid for all times τ, holds for the joint distribution of the reservoir and system entropy contributions to the total trajectory entropy production, although each contribution does not do it separately, as given by Eq. (14). Also a similar DFT holds for the joint distribution of the adiabatic and nonadiabatic entropy contributions to the total trajectory entropy, as given by Eq. (13). For the type of NESS discussed here, two-variable joint PDFs are all that is needed for the corresponding DFTs since there are $M=2$ decompositions of the total trajectory entropy production, like Eq. (9), for this case. We have shown an example with $M=3$ for handling experimental errors in the Hatano-Sasa FT. In any case, in the light of Eqs. (1) and (2), obtaining an adequate minimal M decomposition of the total trajectory entropy production constitutes the cornerstone toward the derivation of generalized FTs for nonequilibrium systems.

This work was supported by CNEA, CONICET (Grant No. PIP11220090100051), and ANPCYT (Grant No. PICT2007886). V.L. was supported in part by the Swiss NSF under MaNEP and Division II, and thanks CNEA for hospitality.

[1] C. Bustamante, J. Liphardt, and F. Ritort, Phys. Today **58**(7), 43 (2005).
[2] F. Ritort, Adv. Chem. Phys. **137**, 31 (2008).
[3] Denis J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[4] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).
[5] J. Kurchan, J. Phys. A **31**, 3719 (1998).
[6] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[7] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997); Phys. Rev. E **56**, 5018 (1997).
[8] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998); Phys. Rev. E **61**, 2361 (2000).
[9] J. Kurchan, J. Stat. Mech.: Theory Exp. (2007), P07005.
[10] R. J. Harris and G. M. Schütz, J. Stat. Mech.: Theory Exp. (2007), P07020.
[11] Y. Oono and M. Paniconi, Prog. Theor. Phys. Suppl. **130**, 29 (1998).
[12] Takahiro Hatano and Shin-ichi Sasa, Phys. Rev. Lett. **86**, 3463 (2001).
[13] T. Speck and U. Seifert, J. Phys. A **38**, L581 (2005).
[14] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[15] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, J. Stat. Mech.: Theory Exp. (2006), P08001.
[16] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).
[17] A similar splitting of the entropy has been considered in [6] but applied for the case of constant σ.
[18] K. Sekimoto, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[19] R. García-García, V. Lecomte, A. B. Kolton, and D. Domínguez (unpublished).
[20] R. van Zon and E. G. D. Cohen, Phys. Rev. Lett. **91**, 110601 (2003).
[21] E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante, and J. Liphardt, Proc. Natl. Acad. Sci. U.S.A. **101**, 15038 (2004).
[22] J. Prost, J.-F. Joanny, and J. M. R. Parrondo, Phys. Rev. Lett. **103**, 090601 (2009).