Estimating Dissipation from Single Stationary Trajectories

Édgar Roldán and Juan M. R. Parrondo

Departamento de Física Atómica, Molecular y Nuclear and GISC, Universidad Complutense de Madrid, 28040-Madrid, Spain

In this Letter we show that the time reversal asymmetry of a stationary time series provides information about the entropy production of the physical mechanism generating the series, even if one ignores any detail of that mechanism. We develop estimators for the entropy production which can detect nonequilibrium processes even when there are no measurable flows in the time series.

The relationship between irreversibility and entropy production forms the core of thermodynamics and statistical mechanics. However, it had not been formulated quantitatively until the recent introduction of the Kullback-Leibler distance or relative entropy in the context of fluctuation and work theorems [1]. The relative entropy between two probability distributions, $p(x)$ and $q(x)$ is defined as

$$D(p\|q) \equiv \sum_x p(x) \log\frac{p(x)}{q(x)}, \qquad (1)$$

and is a measure of their distinguishability [2]. The average entropy production associated with a process driven by an external agent turns out to be equal to the relative entropy between the two probability distributions describing the process running forward and backward in time [1,3–6]. This relative entropy can be thought of as the distinguishability between the process and its time reverse, i.e., as the irreversibility exhibited by the process. The relationship between entropy production and relative entropy has been derived in different scenarios, Hamiltonian dynamics [1,3] and Langevin dynamics [5], and has also been tested in experimental situations [5].

When applied to nonequilibrium stationary states (NESS), the entropy production per unit time reads

$$\frac{\langle \dot{S} \rangle}{k} = \lim_{t\to\infty} \frac{1}{t} D[p(\{x(\tau)\}_{\tau=0}^t)\|p(\{x(t-\tau)\}_{\tau=0}^t)], \qquad (2)$$

where k is the Boltzmann constant and $p(\{x(\tau)\}_{\tau=0}^t)$ is the probability of observing a given trajectory $\{x(\tau)\}_{\tau=0}^t$ in phase space. Since we focus on stationary trajectories—where the external forcing, if any, is constant—there is no need to reverse the driving in the backward process. Moreover, a sufficiently long single trajectory can provide all the necessary statistics to compute the relative entropy in Eq. (2) and consequently the entropy production rate.

Fortunately, the full information of the trajectory in the phase space is not always necessary. Equation (2) follows immediately from the Gallavotti-Cohen theorem [7], by replacing the relative entropy between trajectories with $D(p_S(s)\|p_S(-s))$, where $p_S(s)$ is the probability to observe an entropy production s in a time interval $[0, t]$. In general, the relative entropy calculated using partial information, $\{\tilde{x}(\tau)\}_{\tau=0}^t$ where $\tilde{x}(\tau)$ is a noninvertible function of $x(\tau)$, only provides a lower bound on the average entropy production [1,6,8]. For stationary trajectories, instead of Eq. (2) one obtains a lower bound, which is met if $\tilde{x}(\tau)$ univocally determines the entropy production s.

For discrete stationary trajectories $x_1, \ldots, x_n$, we can define the relative entropy of n-strings as

$$D_n(p_F\|p_B) \equiv \sum_{x_1,\ldots,x_n} p(x_1, \ldots, x_n) \log\frac{p(x_1, \ldots, x_n)}{p(x_n, \ldots, x_1)}. \qquad (3)$$

Following the above arguments, we arrive at

$$\frac{\langle \dot{S} \rangle}{k} \geq d(p_F\|p_B) \equiv \lim_{n\to\infty} \frac{1}{n} D_n(p_F\|p_B). \qquad (4)$$

This equation reveals a striking connection between physics and the statistics of a time series. The left-hand side is a purely physical quantity (it is proportional to the average dissipated energy per step), whereas the right-hand side is a statistical magnitude depending solely on the data $x_1, x_2, \ldots$, but not on the physical mechanism generating those data. Such a connection is a generalization of Landauer's principle relating entropy production and logical irreversibility [1,9,10]. Equation (4) extends this principle and suggests that we can determine the entropy production of an arbitrary NESS by computing the relative entropy of forward and backward trajectories. We could, for instance, determine whether a biological process is active or passive or even estimate, or bound, the amount of consumed ATP by measuring the relative entropy of data generated in the process.

In this Letter we explore the feasibility of such a technique by analyzing the validity of Eq. (4) and developing estimators of the relative entropy. Our approach is general, but we use a discrete flashing ratchet as a case study, wherein direct comparison between analytical and empirical values of the relative entropy and the entropy production is possible. There have been previous attempts to distinguish between equilibrium and NESS. Martin *et al.* checked the fluctuation dissipation relationship in

experimental data from hair bundles of hair cells [11], but this approach needs two types of data: spontaneous and forced fluctuations. Amann *et al.* analyzed the possibility to discriminate between equilibrium and nonequilibrium in a three state chemical system [12]. Finally, Kennel introduced in [13] criteria based on compression algorithms to distinguish between symmetric and asymmetric time series in the context of chaotic signals, without any connection to dissipation. As we show in this Letter, relative entropy provides a more general and simpler framework for the problem of distinguishing between equilibrium and NESS and, moreover, yields estimations and lower bounds on the entropy production.

Two strategies have been considered to estimate the relative entropy between stochastic processes: the first is based on brute-force counting of n-strings, obtaining empirical estimates of $p(x_1, \ldots, x_n)$, and computing D_n using Eq. (3); the second is based on string parsing, the basic procedure of the Lempel-Ziv compression algorithm [14].

The first strategy is simpler and more effective for Markov chains. Our results indicate that this is still the case for some non-Markov process [15]. Consequently, we will restrict ourselves in this Letter to estimations of relative entropy from empirical probability distributions.

If the process and its reverse are Markovian, $p(x_1, x_2, \ldots, x_n) = p(x_1)p(x_2|x_1)\cdots p(x_{n-1}|x_n)$, the relative entropy rate d defined in Eq. (4) can be expressed in terms of the relative entropy between distributions of substrings of size 2:

$$d(p_F\|p_B) = \sum_{x_1,x_2} p(x_1, x_2) \log\frac{p(x_2|x_1)}{p(x_1|x_2)} = D_2 - D_1. \quad (5)$$

In the specific case of a trajectory and its reverse, the one-time statistics are identical and $D_1(p_F\|p_B) = 0$. Then for Markovian dynamics $d(p_F\|p_B) = D_2$, which can be calculated by frequency counting if the number of states and possible transitions is not large. In general, if one defines

$$d_k \equiv D_k - D_{k-1} \quad (6)$$

then $d_k \to d$ for $k \to \infty$. The limit is reached for finite k for the so-called kth order Markov chains, i.e., when blocks of size k, $X_k \equiv (x_n, \ldots, x_{n+k-1})$, are Markovian [16]. In this case $d(p_F\|p_B) = d_{k+1} = d_{k+2} = \cdots$. For more general processes, we will use the following ansatz, proposed in Ref. [17] for Shannon entropy estimation:

$$d_k = d_\infty - c\frac{\log k}{k^\gamma}, \quad (7)$$

where c and γ are parameters that, together with d_∞, can be obtained by fitting the empirical values of d_k vs k.

We have tested the accuracy of these estimators and of the bound (4) in a specific example: a discrete flashing ratchet [18], consisting of a particle moving in a one-dimensional lattice. The particle is at temperature T and moves in a periodic and asymmetric potential of height $2V$,

which is switched on and off at a rate r (see Fig. 1). Trajectories are described by two variables: the position of the particle, $x = \{0, 1, 2\}$, and the state of the potential (on or off), $y = \{0, 1\}$.

To define the dynamics of the particle, we start with a continuous time description based on rates of spatial jumps and switching. We assume that the motion in each potential obeys detailed balance: $k_{i\to j} = e^{-\beta[(V_j - V_i)/2]}$, and $k_{i'\to j'} = 1$ for $i, j = 0, 1, 2$ with $i \neq j$. The system is driven out of equilibrium by imposing constant switching rates $k_{i\to i'} = k_{i'\to i} = r$, $i = 0, 1, 2$, which do not obey detailed balance.

We will focus on the dissipation per step: from the continuous trajectory $[x(t), y(t)]$ we generate a series (x_n, y_n) comprising the states visited by the system. That is, we drop the information of the times when jumps or switches occur. (x_n, y_n) is a Markov chain with transition probabilities given by $p_{\alpha\to\gamma} = k_{\alpha\to\gamma}/\sum_\gamma k_{\alpha\to\gamma}$, with $\alpha, \gamma = 0, 1, 2, 0', 1', 2'$. Introducing these probabilities in Eq. (5), $d(p_F\|p_B) = \beta\sum\langle V_\alpha - V_\gamma\rangle$, where the sum runs over transitions mediated by the thermal bath, $i \to j$, $i' \to j'$. The relative entropy turns out to be the average dissipation per step in units of kT and we recover the main result, Eq. (2) [19]. It is also interesting to explore the relationship between d_2 and the stationary flows $J_{\alpha\gamma} = p_{\alpha\gamma} - p_{\gamma\alpha}$ between states $\alpha, \gamma = 0, 1, 2, 0', 1', 2'$. If $J_{\alpha\gamma} \ll p_{\alpha\gamma}$, we have

$$d_2 \simeq \sum_{\alpha\gamma} \frac{(J_{\alpha\gamma})^2}{2p_{\alpha\gamma}} = \sum_{\alpha<\gamma} \frac{(J_{\alpha\gamma})^2}{p_{\alpha\gamma}}, \quad (8)$$

which is a well-known expression of the entropy production in continuous Markov systems [20], where $d_2 = d$.

Figure 2 shows the dissipation, calculated analytically by solving the six-state Markov chain in the stationary regime, and the estimations discussed above. Because of Markovianity, relative entropies, d_k, immediately converge $d = d_2 = d_3 = \cdots$ and d is equal to the entropy production per step. As long as one has a good estimation of $p(x_1, \ldots, x_k)$, our approach provides accurate values of the entropy production, which is the case for weak potentials $V \simeq kT$. If $V \gg kT$, then uphill jumps, $0 \to 1$, $0 \to 2$, and $1 \to 2$, are so unlikely that they do not occur in a finite trajectory. The higher order the statistics, the earlier this problem arises, as shown in Fig. 2. The reason is that
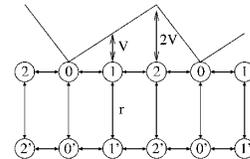

FIG. 1. Discrete ratchet scheme. Particles can jump between the states $i \to j$, $i' \to j'$, and $i \to i'$ in a flashing asymmetric potential of height $2V$ with periodic boundary conditions. The switching rate of the potential is r.
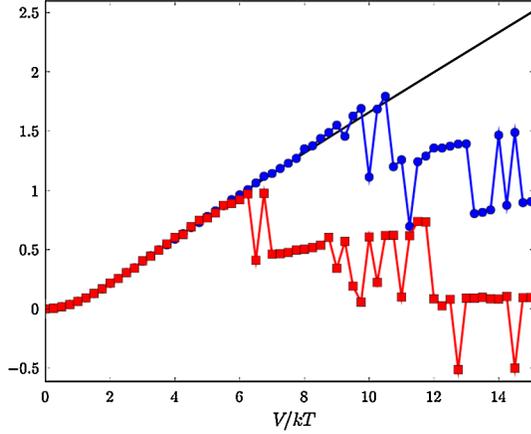
FIG. 2 (color online).   Average dissipation per step (in units of kT) in the flashing ratchet ($r = 1$) and different estimations of relative entropy using a trajectory with $n = 10^6$ steps and full information, as a function of V/kT: analytical calculation of the average dissipation (black line), d_2 (blue circles), d_3 (red squares).

d_3 involves probability distributions of three-step trajectories, the sampling space is bigger, and it is easier that some transitions $i \rightarrow j \rightarrow k$ do not appear while their reverse does. Although these jumps are very unlikely, they contribute significantly to d, as shown in Fig. 2, where d_2 and d_3 have been calculated by restricting the sum in D_k to strings satisfying $p(x_1, \ldots, x_k) \neq 0$ and $p(x_k, \ldots, x_1) \neq 0$.

In real applications, it is more likely that one has only partial information of the trajectories. To study the accuracy of our estimators and of the inequality (4) in this case, we remove the information of the state of the potential and consider trajectories described only by the position $\{x_k\}_{k=1}^n$, which in general are not Markovian. As a consequence, the estimation of the relative entropy $d(p_F \| p_B)$ is more difficult, but even a good estimation of d only provides a lower bound on the relative entropy. It is known that the Gallavotti-Cohen symmetry does not hold in the continuous flashing ratchet if the state of the potential is not considered [21]. In fact, the bound (4) can be quite loose. For instance, if $r \rightarrow \infty$, switching is very fast and the particle moves in an effective potential (the average of on and off) which is periodic. The position x_k becomes Markovian and the current vanishes. Using Eq. (8) one arrives at $d = d_2 = 0$, whereas the dissipation per step is nonzero.

In most cases, however, the bound given by Eq. (4) provides significant information. In Fig. 3 we show the estimation of d using the empirical values of d_k for $k = 2, 9$, and the extrapolation d_∞ resulting from the fit of the ansatz in Eq. (7). The error bars in Fig. 3 correspond to the error in the fit with a confidence interval of 90%. Our estimations clearly distinguish between the equilibrium case ($V = 0$) and the NESS. The empirical d_k with $k > 3$ correctly reproduce the order of magnitude of the actual
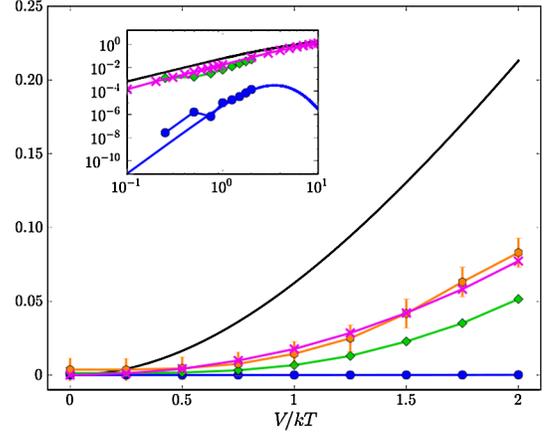

FIG. 3 (color online).   Average dissipation per step (in units of kT) in the flashing ratchet ($r = 1$) and different estimations of relative entropy using a trajectory with $n = 10^7$ steps and partial information (position) as a function of V/kT: analytical calculation of the average dissipation (black line), d_2 (blue circles), d_9 (green diamonds), d_∞ in Eq. (7) (orange hexagons with error bars), and Monte Carlo semianalytical calculation of d (purple crosses). Inset: Estimators for weak potentials in a log-log plot. We have added in the inset the analytical calculation of d_2 (blue solid line).

dissipation (see inset in Fig. 3), although they underestimate it. There are two possible causes for this deviation: either we are underestimating the actual relative entropy d or the bound provided by Eq. (4) is not tight. To clarify this question we need an analytical calculation of the relative entropy between two non-Markov processes. In our case, the relative entropy D_n reads

$$D_n = \left\langle \log \frac{\sum_{y_1, \ldots, y_n} p(x_1, y_1; \ldots; x_n, y_n)}{\sum_{y_1, \ldots, y_n} p(x_n, y_n; \ldots; x_1, y_1)} \right\rangle, \qquad (9)$$

where the average is taken over all possible trajectories. The probability distribution $p(x_1, y_1; \ldots; x_n, y_n) = p(x_1, y_1) \times p(x_2, y_2 | x_1, y_1) \times \cdots \times p(x_n, y_n | x_{n-1}, y_{n-1})$ is known, but Eq. (9) cannot be calculated exactly. Fortunately, the log in Eq. (9) is a self-averaging quantity for large n [22] and we can compute the average using a single long typical trajectory [15]. We show in Fig. 3 the value of d obtained by this Monte Carlo semianalytical calculation (purple crosses), which is very close to the estimation d_∞ based on the ansatz Eq. (7).

Although the relative entropy d underestimates the actual dissipation, it does reproduce its asymptotic behavior. Entropy production decreases as V^2 for small V, so do d_∞ and d_9 (see inset of Fig. 3). On the other hand, $d_2 \propto V^6$, since the current is $J \propto V^3$ [see Eq. (8)].

We have found in several instances a similar qualitative improvement on the estimation of relative entropy when using blocks of size bigger than two. In particular, d_3 and above outperform d_2, which, as indicated by Eq. (8), is
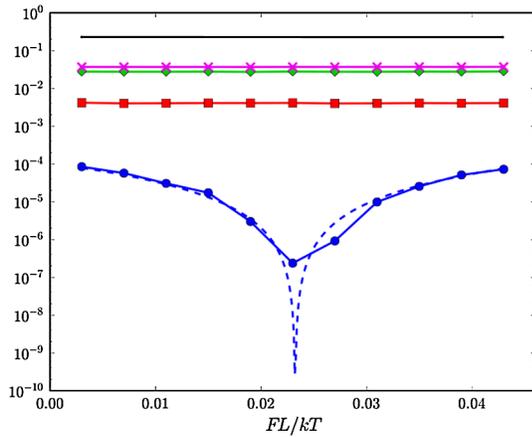
FIG. 4 (color online).   Average dissipation per step (in units of kT) in the flashing ratchet ($r = 2$, $V = 2kT$) with external force F and different estimations of relative entropy using a trajectory with $n = 10^7$ steps with partial information (position): analytical calculation of the average dissipation (black line), d_2 (blue circles, analytical values in blue dashed line), d_3 (red squares), d_9 (green diamonds), and semianalytical calculation of d (purple crosses). The minimum in d_2 corresponds to the stall force.

equivalent to the standard calculation of entropy production using the currents observable from the available data, in our case, the spatial current. For a striking illustration of this effect we add an external force F to the flashing ratchet and study dissipation and relative entropy close to the stalling force F_{stall}, for which the spatial current and d_2 both vanish. Jumping rates are now biased in the direction of the force, giving the following detailed balance condition $k_{i \to j}/k_{j \to i} = e^{-\beta(V_j - V_i - FL_{ij})}$, $L_{ij} = 1$ being the distance between i and j.

We have plotted in Fig. 4 the real dissipation, the analytical value of d and d_2, and the empirical values of d_2, d_3, and d_9, close to the stalling force F_{stall}. Recall that, for $F = F_{\text{stall}}$, the position of the particle does not exhibit any flow and its average position remains constant. Consequently, d_2 or any other estimation of entropy production based on flows will fail. However, the relative entropy calculated using blocks of size 3 captures the nonequilibrium nature of the time series.

In conclusion, we have shown that the statistical properties of a time series impose a lower bound on the entropy produced in generating the series. This lower bound is valid even if we do not have any access or information of the physical mechanism generating the data. Finally, we have shown that the bound can be nontrivial, predicting dissipation even when the data do not exhibit any measurable flow. Our techniques could be applied to data from different sources. In the case of biological systems, they could help to distinguish between passive and active processes, and even to estimate ATP consumption. On the other side, as in the case of Landauer's principle, relative entropy can be used to ascertain the minimal entropy production associated with a specific behavior, such as spatiotemporal patterns, excitable systems, etc. This in turn may influence the design of optimal devices with functionalities given by these behaviors.

[1] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007).
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, Hoboken, NJ, 2006), 2nd ed.
[3] J. M. R. Parrondo, C. V. den Broeck, and R. Kawai, New J. Phys. **11**, 073008 (2009).
[4] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[5] D. Andrieux, P. Gaspard, S. Ciliberto, N. Garnier, S. Joubaud, and A. Petrosyan, Phys. Rev. Lett. **98**, 150601 (2007).
[6] J. Horowitz and C. Jarzynski, Phys. Rev. E **79**, 021106 (2009).
[7] E. Cohen and G. Gallavotti, J. Stat. Phys. **96**, 1343 (1999).
[8] A. Gomez-Marin, J. M. R. Parrondo, and C. V. den Broeck, Europhys. Lett. **82**, 50 002 (2008).
[9] R. Landauer, IBM J. Res. Dev. **44**, 261 (2000).
[10] D. Andrieux and P. Gaspard, Proc. Natl. Acad. Sci. U.S.A. **105**, 9516 (2008).
[11] P. Martin, A. J. Hudspeth, and F. Jülicher, Proc. Natl. Acad. Sci. U.S.A. **98**, 14 380 (2001).
[12] C. P. Amann, T. Schmiedl, and U. Seifert, J. Chem. Phys. **132**, 041102 (2010).
[13] M. B. Kennel, Phys. Rev. E **69**, 056208 (2004).
[14] J. Ziv and N. Merhav, IEEE Trans. Inf. Theory **39**, 1270 (1993).
[15] E. Roldán and J. M. R. Parrondo (to be published).
[16] Z. Rached, F. Alajaji, and L. L. Campbell, IEEE Trans. Inf. Theory **50**, 917 (2004).
[17] T. Schurmann and P. Grassberger, Chaos **6**, 414 (1996).
[18] A. Ajdari and J. Prost, C.R. Acad. Sci. Ser. Gen., Ser. 2 **315**, 1635 (1992).
[19] Each transition $\alpha \to \gamma$ obeying detailed balance contributes to d_2 as $\beta\langle V_\alpha - V_\gamma \rangle$, which is the average entropy increase in the bath due to the transition. This still applies for systems in contact with several baths at different temperatures. In our case, the constant flashing rate r can be interpreted as a transition mediated by a bath at infinite temperature $\beta = 0$, whose entropy does not change when absorbing a finite amount of energy.
[20] J. M. R. Parrondo and B. J. de Cisneros, Appl. Phys. A **75**, 179 (2002).
[21] D. Lacoste and K. Mallick, Phys. Rev. E **80**, 021923 (2009).
[22] P. Jacquet, G. Seroussi, and W. Szpankowski, Theor. Comput. Sci. **395**, 203 (2008).