Thermodynamic cost of acquiring information

Kaonan Micadei* and Roberto M. Serra†
Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, R. Santa Adélia 166, 09210-170 Santo André, São Paulo, Brazil

Lucas C. Céleri‡
Instituto de Física, Universidade Federal de Goiás, 74.001-970, Goiânia, Goiás, Brazil
(Received 16 July 2013; published 12 December 2013)

Connections between information theory and thermodynamics have proven to be very useful to establish bounding limits for physical processes. Ideas such as Landauer's erasure principle and information-assisted work extraction have greatly contributed not only to broadening our understanding about the fundamental limits imposed by nature, but also paving the way for practical implementations of information-processing devices. The intricate information-thermodynamics relation also entails a fundamental limit on parameter estimation, establishing a thermodynamic cost for information acquisition. We show that the amount of information that can be encoded in a physical system by means of a unitary process is limited by the dissipated work during the implementation of the process. This includes a thermodynamic tradeoff for information acquisition. Likewise, the information acquisition process is ultimately limited by the second law of thermodynamics. This tradeoff for information acquisition may find applications in several areas of knowledge.

DOI: 10.1103/PhysRevE.88.062123 PACS number(s): 05.70.Ln, 03.67.Ac, 06.20.−f

I. INTRODUCTION

Information theory first encountered thermodynamics when Maxwell introduced his famous Demon [1]. This relation became clear with Brillouin's treatment of information entropy (due to Shannon) and thermodynamic entropy (due to Boltzmann) on an equal footing [2]. Many advances linking these two apparently distinct areas have been achieved since then, with one of the most remarkable being ascribed to Landauer's erasure principle [3]. This principle, introduced as an effective way to *exorcize* Maxwell's Demon, states that the erasure of information is a logically irreversible process that must dissipate energy. More recently, developments in this direction have included theoretical and experimental investigations of Landauer's principle and its consequences [4,5], work extraction by feedback control of microscopic systems [6–10], and links between the second law of thermodynamics and two fundamental quantum-mechanical principles, i.e., the wave-function collapse [11] and the uncertainty relation [12]. Here, we introduce a thermodynamic tradeoff for information acquisition, which relates the uncertainty of the information acquired in a parameter estimation process with the dissipated work by the encoding process. This tradeoff relation is obtained by making a formal connection between an elusive quantity from estimation theory, called Fisher information [13–16], and the Jarzynski equality [17].

II. THERMODYNAMIC TRADEOFF

Natural sciences are based on experimental and phenomenological facts. Parameter estimation protocols have a central role to observe new phenomena or to validate some theoretical prediction. Suppose we want to determine the value of some parameter, say φ. This can be accomplished, generally, by employing a probe, ρ_T. We will assume that the probe state is initially in thermal equilibrium at absolute temperature T, so ρ_T is the canonical equilibrium (Gibbs) state [18]. To extract some information about the parameter φ, the probe could be prepared (through a process Γ) in a suitable *blank* state represented by ρ_0. Then the probe is transformed by a unitary process Λ_φ in order to encode information about the parameter on the probe state ρ_φ. In general, in real-world applications, these operations are logically irreversible (due to unavoidable fluctuations during the system's dynamics) and therefore must have a thermodynamic cost. The effectiveness of the estimation (metrology) process depends on how information is encoded in the probe system. This encoding operation consumes some work from a thermodynamic point of view. An estimation of the parameter φ can be obtained by a suitable read-out of the encoded probe system ρ_φ. The aforementioned protocol [also outlined in Fig. 1(a)] abstractly summarizes the operation of almost all high-precision measurement devices. Employing this general framework, we show that the uncertainty (the mean-square root error) $\Delta\varphi$ of an estimation process is limited by a general physical principle,

$$\Delta\mathcal{C}\,\Delta\mathcal{I}_\varphi \geqslant \frac{k_B}{2}, \tag{1}$$

where k_B is Boltzmann's constant, and the thermodynamic tradeoff for information acquisition is defined as the mean dissipated work (i.e., work not converted into encoding) $\langle\mathcal{W}_D\rangle$ at a given temperature T as $\Delta\mathcal{C} = \langle\mathcal{W}_D\rangle/T$ and the relative acquired information as $\Delta\mathcal{I}_\varphi = (\Delta\varphi)^2/\delta_\varphi^2$. δ_φ is a quantity describing the accuracy of the encoding process. Roughly speaking, δ_φ is the precision of the experimental device used to implement Λ_φ [the minimum scale for φ; see Fig. 1(b) and the Appendixes]. The symbol $\langle\cdots\rangle$ represents the mean value with respect to an ensemble of measurements. The physical quantities appearing in Eq. (1) are highly process-dependent and must be carefully defined in each physical setup. We

*kaonan.bueno@ufabc.edu.br
†serra@ufabc.edu.br
‡lucas@chibebe.org

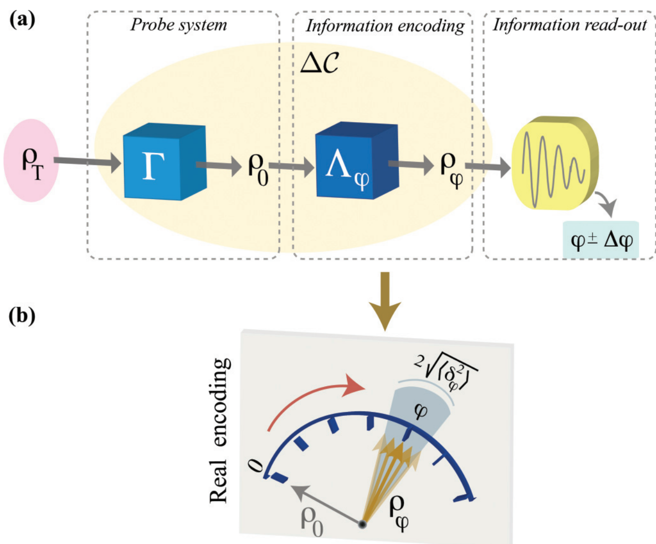

FIG. 1. (Color online) General sketch of a parameter estimation process. (a) The estimation process works as follows: A probe system initially in the thermal equilibrium, ρ_T, at a given temperature T, is prepared in a suitable blank state ρ_0 (through a process Γ). Then, Λ_φ acts on the probe in order to encode some information about the desired parameter φ. The thermodynamic cost of the probe state preparation and encoding of information is $\Delta\mathcal{C}$. The final step is the read-out of the encoded probe ρ_φ. This measurement process results in a probability distribution, p_φ, that contains the information about φ. After some statistical manipulation of this measured distribution, an estimation for the value of φ is obtained with the mean-square root error $\Delta\varphi$. (b) Pictorial representation of an imperfect encoding of information. The probe state is represented as a pointer and the encoding process Λ_φ has a (minimum) finite precision δ_φ.

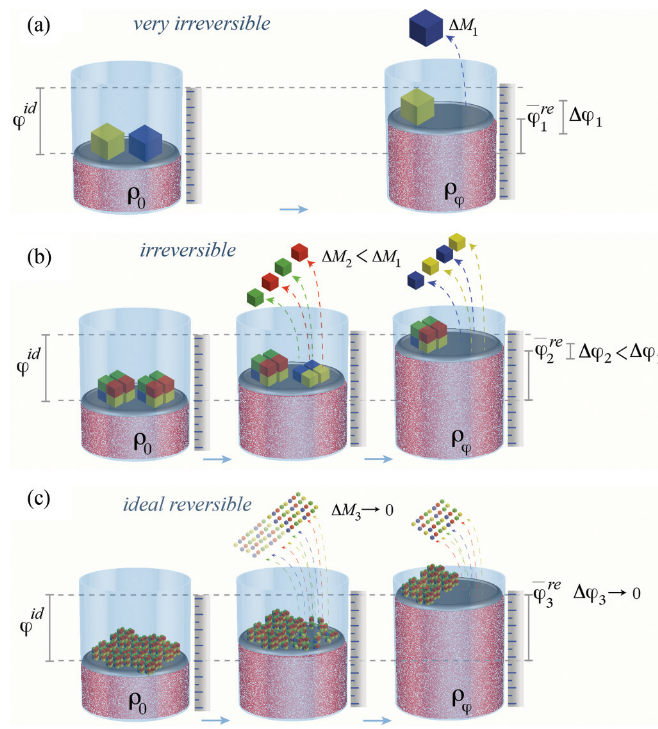

FIG. 2. (Color online) Illustration of a parameter estimation process employing a classical apparatus. The apparatus itself is composed of a gas confined within a cylindrical chamber with a movable piston. Information is encoded in the piston position adding or removing some amount of mass above the piston. In the ideal encoding process (reversible), the position displacement is given by $\varphi^{\rm id}$. On the other hand, in a finite step (irreversible, nonideal) protocol the mean position displacement is labeled as $\bar{\varphi}^{\rm re}$. (a) A very irreversible process in which the information encoding is performed moving suddenly a large amount of mass in a single shot resulting in the displacement $\bar{\varphi}_1^{\rm re}$ with a large uncertainty $\Delta\varphi_1$. (b) In this case, the information encoding is performed moving small portions of mass. This is still an irreversible process which dissipates less work than the first one, leading to the displacement $\bar{\varphi}_2^{\rm re}$ with an uncertainty $\Delta\varphi_2$ (which is smaller than $\Delta\varphi_1$). (c) Sketch of an idealized reversible process in which the amount of mass above the piston is removed in an adiabatic way, in this case $\bar{\varphi}_3^{\rm re} \rightarrow \varphi^{\rm id}$ and $\Delta\varphi_3 \rightarrow 0$.

proceed by analyzing the physical meaning of Eq. (1) and discussing some of its implications, postponing its derivation.

The work consumed in the parameter estimation process could ultimately be attributed to the logical irreversibility of information encoding. The first step for any estimation protocol is to prepare the probe. If we employ an out-of-equilibrium probe, we have to erase the thermal state to prepare the probe in a suitable blank state, and it has some energetic cost (from Landauer's principle). The second step, i.e., the encoding of information in the probe state (Λ_φ), in a realistic apparatus is not perfect. In other words, the finite precision δ_φ of the encoding operation implies logical irreversibility, and, as a consequence, some work is dissipated by fluctuations. For the sake of clarity, let us discuss these issues in two physical contexts.

First, we consider a simple classical thermodynamic system. In this classical setting, all the quantities appearing in Eq. (1) are naturally defined. Nevertheless, our results can be applied to both (out-of-equilibrium) classical and quantum systems. Let us suppose that our apparatus is composed of a gas confined within a cylindrical chamber in which the upper base is made out of a movable piston with some amount of matter (M) placed over it as sketched in Fig. 2. The gas is our probe system and the position of the movable piston indicates the probe state. Defining φ_0 as the equilibrium position of the piston, we can encode information on this system introducing (or removing) some amount of matter over the piston. The information is encoded by the displacement φ of the piston from its initial equilibrium position φ_0. Considering the parameter estimation processes as described in Fig. 1(a), we have the following steps: (i) Initially, the system is in thermal equilibrium at temperature T, being described by the state ρ_T, corresponding to the piston position φ_0 (with some amount of mass M on the piston). In this situation, all the forces acting on the piston are in equilibrium and its position is fixed (except for thermal fluctuations). In this context, the equilibrium state ρ_T could be a suitable probe ($\rho_0 = \rho_T$), so the probe preparation process [Γ of Fig. 1(a)] does not dissipate work. (ii) By removing (in an adiabatic or nonadiabatic way) some amount of mass (ΔM) on the piston, the information can be encoded into the probe. Due to the imbalance of forces, this operation drives the piston to a new equilibrium position described by the state ρ_φ, corresponding to position φ, thus encoding the information into the probe state. During such an implementation, a certain amount of work must be employed and part of it may be dissipated into the environment. To get a good estimation of the encoded information, this protocol should be repeated N times, or N identical copies of such a system should be employed. Each realization of the protocol is driven by an amount of performed work w^i. The mean applied work is then given by $\langle \mathcal{W} \rangle = \sum p(w^i) w^i$, with $p(w^i)$ being the measured work probability distribution for the whole ensemble of realizations. (iii) Measuring the piston position in all realizations, we obtain an estimation $\varphi = \langle \varphi \rangle \pm \Delta\varphi$ for the parameter, where $\langle \varphi \rangle = \sum p(\varphi_i)\varphi_i$ is the mean value of the piston displacement, $p(\varphi_i)$ is the observed probability distribution, and $\Delta\varphi$ is the mean-square root deviation.

In the example explored above, the work is dissipated in step (ii). When we remove a certain amount ΔM of mass from the top of the piston, the gas expands until a new equilibrium position is reached. The amount of removed mass determines how much the piston position changes, and, ultimately, it will also determine how much work will be dissipated. If the whole mass is removed in just one shot [Fig. 2(a)], there will be a huge amount of dissipated work, since the gas will expand from the initial state (corresponding to ρ_T) to the final one (corresponding to ρ_φ) through a sudden and irreversible path [19]. The amount of dissipated work is given by $\langle \mathcal{W}_D \rangle = \langle \mathcal{W} \rangle - \Delta F$, where ΔF is the difference in free energy between the final and the initial states, and $\langle \mathcal{W} \rangle$ is the mean invested work during the encoding process. This example clearly shows that some information is lost in the encoding process due to work dissipation caused by finite changes in the system (irreversibility).

To minimize the information loss and, consequently, to improve the precision of the protocol, we have to diminish fluctuations during the dynamics as much as possible. This can be accomplished by removing the mass in small portions, as depicted in Fig. 2(b), with the limit being the idealized reversible process for which $\langle \mathcal{W}_D \rangle = 0$ [Fig. 2(c)]. In this case, Eq. (1) seems to be flawed, but a deeper analysis reveals that this is not the case. For the implementation of a reversible process, we must take the limit $\Delta M \to 0$ (the process must be implemented in a quasistatic way). But, this limit implies $\delta_\varphi \to 0$, since δ_φ (the encoding accuracy) is the minimum step size that the piston is able to move, i.e., the minimum change in the system (see below and the Appendixes for formal details). In the limit of reversible processes, we can read out all the information encoded in the probe. On the other hand, in the real word the "scale" of the encoding apparatus is finite. In this case, the minimum amount of mass that can be removed is finite; this also introduces a minimum step size for the position of the piston, i.e., $\delta_\varphi > 0$. This inevitable fact leads to information loss in the encoding process, implying some work dissipation. In fact, any realistic encoding apparatus with finite precision (scale) is irreversible (in an ensemble of realizations), therefore the apparatus must dissipate work, introducing uncertainty in the parameter estimation as ultimately bounded by Eq. (1).

An important point concerning the analysis of Eq. (1) is the relation between the dissipated work $\langle \mathcal{W}_D \rangle$ and the accuracy of the encoding process δ_φ. As illustrated in the above example and in Fig. 2, in order to get $\langle \mathcal{W}_D \rangle \to 0$ in an ensemble of realizations, we have to employ an apparatus with almost perfect accuracy (adiabatic encoding), $\delta_\varphi \to 0$. Nevertheless, all real processes are irreversible and information is inevitably lost, since we have a finite step size in a real encoding process ($\delta_\varphi > 0$) implying finite dissipated work ($\langle \mathcal{W}_D \rangle > 0$). The physical limit on the variance ($\Delta_\varphi > 0$) is given by the second law, by means of the dissipated work in the encoding process. This amount of work is a direct consequence of the finite precision $\delta_\varphi > 0$ in every real encoding process. In other words, due to the finite encoding precision, some amount of information is lost, leading to the dissipated work. The saturation of Eq. (1) occurs not in the limit $\langle \mathcal{W}_D \rangle \to 0$ (which would imply $\delta_\varphi \to 0$) but when $\delta_\varphi = \Delta_\varphi$. This can be achieved, for example, if the same apparatus is used in the decoding process with no additional error source. Although we have explored a specific example, the discussion above is independent of the physical system, just as irreversible processes (associated with finite changes in the system) must increase entropy.

Two other important limits are zero and infinite temperature. Let us assume that both dissipated work $\langle \mathcal{W}_D \rangle$ and the encoding accuracy δ_φ are constants with respect to the temperature. When $T \to \infty$, Eq. (1) leads to $(\Delta\varphi)^2 \to \infty$. The observer cannot obtain any information encoded by the process Λ_φ. Actually, we cannot encode any information in this limit due to the infinite amplitude of thermal fluctuations, which wash out all the information, no matter how precise the encoding process is. In the opposite limit, $T \to 0$, we have $(\Delta\varphi)^2 \geqslant 0$. For classical systems, this is a valid limit and the inequality could be, in principle, saturated. However, due to the third law of thermodynamics, it is not allowed for quantum systems to reach this limit. $(\Delta\varphi)^2$ is always greater than zero due to quantum fluctuations.

The bound presented in Eq. (1) also holds for quantum strategies for parameter estimation employing out-of-equilibrium probes. Now, let us consider a standard interferometric strategy to estimate a phase shift between two states. This task can be accomplished by observing the probability for the measurement of the probe in a suitable basis. We are going to label the two states by $|0\rangle$ and $|1\rangle$. A suitable probe in this case is a balanced superposition, $|\psi_0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ ($\rho_0 = |\psi_0\rangle\langle\psi_0|$). This probe state preparation process could be a suitable postselected projective measurement on the thermal equilibrium state ρ_T. The encoding of information could be carried out by a phase shifter, such as, for example, $U(\varphi) = e^{i\phi|1\rangle\langle 1|}$. In a real interferometer, the minimum step size δ_φ to encode the phase φ on the probe state is finite. Therefore, we have an imperfect encoding [as pictorially described in Fig. 1(b)], and in an ensemble of realizations, the evolution is irreversible. This finite accuracy of the phase shifter ($\delta_\varphi > 0$) implies information loss and work consumption in the encoding process Λ_φ. In fact, the probe state preparation is also a nonideal process introducing another source of dissipated work. The bound for information acquisition in the out-of-equilibrium quantum context is also given by Eq. (1).

III. OUTLINE OF THE DERIVATION

Now, we outline the derivation of Eq. (1). Consider again the general framework for a parameter estimation process described in Fig. 1(a). To inspect how a given unbiased estimator for the parameter φ is close to the real encoded information, we can use the so-called Cramér-Rao bound [20,21],

$$(\Delta\varphi)^2 \geqslant \frac{1}{\mathcal{F}}, \tag{2}$$

where $\mathcal{F}$ is the Fisher information, usually defined as $\mathcal{F} = \int dx\, p_\varphi(x)(\frac{\partial \ln p_\varphi(x)}{\partial\varphi})^2$. $p_\varphi(x)$ is the probability distribution for the best read-out strategy of the encoded probe, and it contains the information about φ. For our proposal, it will be interesting to express the Fisher information in terms of relative entropy as [15,16]

$$\mathcal{F} \approx 2\frac{S\big(p_\varphi^{\rm re}||p_\varphi^{\rm id}\big)}{\delta_\varphi^2}, \tag{3}$$

with $p_\varphi^{\rm re}$ being the read-out probability distribution obtained in a real (irreversible, nonideal) experiment, $p_\varphi^{\rm id}$ being the read-out probability distribution of an ideal (reversible) parameter estimation protocol, and δ_φ is the accuracy of the process. Here, we refer to the ideal process as the limit process of the second law, i.e., the adiabatic process where all the work in converted into free energy ($\langle\mathcal{W}\rangle - \Delta F = 0$). By the real (irreversible) experiment we mean that the parameter estimation apparatus is nonideal in the sense that the process running in such an apparatus will dissipate some amount of work ($\langle\mathcal{W}\rangle - \Delta F > 0$) due to fluctuations. In this case, we consider a slight nonideal process working very near to the reversible limit, since we are interested in a high-precision measurement apparatus. The approximation presented in Eq. (3) is a very good approximation in this setting (see Appendix C for details).

Next, we relate the Fisher information with information loss and the dissipated work through a formal relation between work and information (obtained in Appendix A). Considering that a system is driven through the injection of work $\mathcal{W}$ by an external agent from the initial equilibrium state to some final one, Jarzynski proved that [17]

$$\langle e^{-\mathcal{W}/k_BT}\rangle = e^{-\Delta F/k_BT}, \tag{4}$$

where the mean is computed over the ensemble of realizations and ΔF is the free-energy difference between the final and the initial system's states. T is the temperature of the initial equilibrium state. In Appendix A, we obtain the following information-theoretic relation:

$$\frac{\langle\mathcal{W}\rangle - \Delta F}{k_BT} = \langle\mathcal{I}_{x,x_\varphi}\rangle, \tag{5}$$

where $\langle\mathcal{I}_{x,x_\varphi}\rangle$ is the mutual information between the readout distribution for the final probe state (where the parameter φ is encoded, ρ_φ) and the distribution for the initial thermal state (ρ_T). x is some parameter characterizing the distribution. It is possible to obtain the Jarzynski equality from this development, as shown by Vedral [27] (see also Appendix A).

In addition to the above results, we can show that

$$\langle\mathcal{I}_{x,x_\varphi}\rangle \geqslant \frac{\delta_\varphi^2}{2}\mathcal{F}, \tag{6}$$

where $\mathcal{F}$ is the Fisher information (classical or quantum) for the encoded state ρ_φ and

$$\left(\frac{\delta_\varphi}{\varphi^{\rm id}}\right)^2 = \left(\frac{\varphi^{\rm re}}{\varphi^{\rm id}} - 1\right)^2 \ll 1 \tag{7}$$

is the relative accuracy of the estimation process. The approximation in Eq. (7) means that the error in the measurement must be much smaller than the parameter being measured. This is quite a reasonable assumption since an error of the same order of the parameter would render meaningless the entire parameter estimation process.

Combining everything together, from Eqs. (2), (5), and (6), we show that

$$(\Delta\varphi)^2 \geqslant \frac{\delta_\varphi^2}{2\langle\mathcal{I}_{x,x_\varphi}\rangle}, \tag{8}$$

which is our main result expressed in Eq. (1) as a tradeoff relation.

IV. DISCUSSIONS

We introduced a physical principle that bounds information acquisition, Eq. (1), derived from an information-theoretic relation, connecting the Jarzynski equality [17] with the Cramér-Rao [20,21] bound. This is a general result, applicable to classical or quantum contexts, stating that the amount of information that can be encoded by means of a unitary process is limited by the dissipated work (due to logical irreversibility) during the implementation of the estimation process. This conclusion reveals a deep connection between metrology and thermodynamics, implying that the physical limit for the precision of a parameter estimation process (which is equivalent to encoding and decoding information processes) is given by thermodynamics. Moreover, the lower bound on the uncertainty about the estimation of a given parameter is zero only in the thermodynamic limit of reversible (adiabatic) processes (imposed by the second law).

Inequality (1) could be conceived as a counterpart of Landauer's principle, as both of them are assertions about the work cost of information (acquisition or erasure). Furthermore, it would be interesting to investigate the relationship of the results herein with generalized uncertainty relations. At this point, it is reasonable to presume that the basic principles of quantum mechanics itself are probably subtly connected to the second law of thermodynamics [12,22,23] in an informational scenario.

From the point of view of the experimental verification of Eq. (1), it is important to precisely establish the system in order to define all the quantities involved, such as the work employed in the process and how the information is encoded and read out. Discussing the fundamentals of physics, Planck has argued that the number of dimensional fundamental constants in nature should be equal to four [24]: the Newtonian gravitational constant G, the speed of light c, and Planck's and Boltzmann's constants h and k_B, respectively. The authors of Ref. [25] concluded that this number should be two, chosen between G, c, and h, having discarded Boltzmann's constant for being a conversion factor between temperature and energy. In Ref. [26], the viewpoint that Planck's constant is superfluous was advocated and k_B was also discarded for the same reason as given in [25]. If we define temperature as twice the mean value of the energy stored in each degree of freedom of a system in thermal equilibrium, $T = 2\langle E_0\rangle$, k_B turns out to be a dimensionless quantity equal to 1 and Eq. (1) becomes

$$\frac{\langle\mathcal{W}_D\rangle}{\langle E_0\rangle}\left(\frac{\Delta\varphi}{\delta_\varphi}\right)^2 \geqslant 1, \tag{9}$$

which means that the precision of the information acquired in a parameter estimation process is limited by the mean dissipated work per degree of freedom of the encoding system. From a more practical standpoint, inequality (1) is quite meaningful for technological applications in metrology relating the reversibility of a high-precision measurement device with its efficiency.

ACKNOWLEDGMENTS

We warmly acknowledge K. Modi for insightful discussions and useful comments. This work was supported by FAPESP, CAPES, CNPq, and the Brazilian National Institute for Science and Technology of Quantum Information (INCT-IQ).

APPENDIX A: INFORMATION-WORK RELATION

In this Appendix, we will obtain relation (5) of the main text. Let us consider a quantum system externally driven through some process such that its initial Hamiltonian is described by $H = \sum_n E_n |\psi_n\rangle\langle\psi_n|$ (in the spectral basis) and its final Hamiltonian reads $H' = \sum_m E'_m |\psi'_m\rangle\langle\psi'_m|$. The initial system state is taken as ρ_0 and the probability distribution for the occupation of the initial Hamiltonian eigenstates is given by $p(n) = \mathrm{Tr}\rho_0 |\psi_n\rangle\langle\psi_n|$. Considering that the system evolves through some process to a final state ρ', we obtain a distribution $p(m) = \mathrm{Tr}\rho' |\psi'_m\rangle\langle\psi'_m|$ for the occupation of the final Hamiltonian eigenstates. A key quantity in our derivation is the mutual information between the joint probability distribution of the outcomes in the measurements of the initial and final Hamiltonian eigenstates, $p(m,n)$. This mutual information can be obtained from the information density, $\mathcal{I}_{n,m} \equiv \ln[p(n,m)/p(n)p(m)]$, as

$$\langle \mathcal{I}_{n,m}\rangle = \sum_{m,n} p(m,n) \ln \frac{p(m,n)}{p(m)p(n)}. \tag{A1}$$

The joint probability distribution $p(m,n)$, which can be written as $p(m,n) = p(m|n)p(n)$ [where $p(m|n)$ corresponds to the transition probability from the initial state n to the final state m], is not directly observable [28]. On the other hand, we have direct access to the initial, $p(n)$, and final, $p(m)$, distributions. In a microscopic thermodynamics description, it is possible to reach the Gibbs ensemble from a distribution which maximizes the Shannon entropy—this means a maximization over the fluctuations—satisfying normalization and thermal energy constraints [29]. Employing the same reasoning, we will find the distribution $p(m,n)$ which provides us with the maximal information $\langle I_{n,m}\rangle$ over the fluctuations, during the process that drives the system Hamiltonian from H to H', with the following constraints:

$$\sum_{m,n} p(m,n) = 1, \tag{A2}$$

$$\sum_{m,n} (E'_m - E_n) p(m,n) = \langle H'\rangle - \langle H\rangle = \langle \Delta E\rangle, \tag{A3}$$

where $\Delta E \equiv E'_m - E_n$, the first condition is the normalization, and the second one is an energy constraint in such a way that the distribution $p(m,n)$ gives us the mean energy variation during the process. This approach is justified further, since it implies known results of fluctuation theorems such as the Jarzinski equality.

Assorting Lagrange multipliers λ_0 and λ_1 to the first and second constraints, we have

$$p(k,l)\left[\sum_{m,n} p(m,n) \ln \frac{p(m,n)}{p(m)p(n)} + \lambda_0 \left(\sum_{m,n} p(m,n) - 1\right) + \lambda_1 \left(\sum_{m,n} (E'_m - E_n) p(m,n) - \langle \Delta E\rangle\right)\right] = 0. \tag{A4}$$

Since the variations on the probability distribution elements are independent, Eq. (A4) is satisfied if

$$\ln \frac{p(k,l)}{p(k)p(l)} + 1 + \lambda_0 + \lambda_1 (E'_k - E_l) = 0.$$

Therefore, $p(k,l) = p(k)p(l)e^{-1-\lambda_0-\lambda_1(E'_k - E_l)}$. From the normalization constraint introduced in Eq. (A2), it follows that

$$\sum_{m,n} p(m)p(n)e^{-1-\lambda_0-\lambda_1(E'_m - E_n)} = e^{-1-\lambda_0} \sum_{m,n} p(m)p(n)e^{-\lambda_1(E'_m - E_n)} = 1, \tag{A5}$$

which implies $e^{1+\lambda_0} = \mathcal{Z}$, where we have defined $\mathcal{Z} \equiv \sum_{m,n} p(m)p(n)e^{-\lambda_1(E'_m - E_n)}$. In this way, we can rewrite the joint probability distribution of the initial and final outcomes as

$$p(m,n) = p(m)p(n)\frac{1}{\mathcal{Z}} e^{-\lambda_1(E'_m - E_n)}. \tag{A6}$$

The conditional probability for the occurrence of outcome m in a measurement, on ρ', of the final Hamiltonian eigenstates, H', given that the initial outcome was n, is given by $p(m|n) = p(n,m)/p(n)$, which, from the above relations, turns out to be $p(m|n) = e^{-\lambda_1(E'_m - E_n)} p(m)/\mathcal{Z}$ [$\sum_m p(m|n) = 1$].

From the energy constraint in Eq. (A3), we obtain

$$\sum_{m,n} E'_m p(m,n) = \sum_m E'_m p(m) = \langle H'\rangle,$$
$$\sum_{m,n} E_n p(m,n) = \sum_n E_n p(n) = \langle H\rangle.$$

We note that $p(n)$ is independent of λ_1, so $\langle H\rangle$ does not fix the λ_1 value, since λ_1 is taken as an arbitrary constant expressed in the inverse of energy unit. For any finite λ_1 (with $|\lambda_1| < \infty$), we can use Eq. (A6) to write

$$\sum_{m,n} (E'_m - E_n) p(m)p(n) \frac{1}{\mathcal{Z}} e^{-\lambda_1(E'_m - E_n)} = \langle \Delta E\rangle. \tag{A7}$$

It is easy to see that $-\lambda_1 = \langle \Delta E\rangle$. Now, let us employ the above relation to rewrite the mutual information between the outcomes in the measurements of the initial and final Hamiltonian eigenstates as

$$\langle \mathcal{I}_{n,m}\rangle = -\ln \mathcal{Z} - \lambda_1 \langle \Delta E\rangle. \tag{A8}$$

Taking the variation of $\langle \mathcal{I}_{n,m}\rangle$ relative to $\langle \Delta E\rangle$, we have

$$\langle \Delta E\rangle = -\lambda_1 \langle \Delta E\rangle - \langle \Delta E\rangle\langle \Delta E\rangle - \lambda_1 = -\lambda_1. \tag{A9}$$

Since $\lambda_1(E'_m - E_n)$ must be dimensionless and $\langle \mathcal{I}_{n,m} \rangle$ is a kind of entropy variation, we can assume $-\lambda_1 = \beta = 1/k_B T$ (where k_B is the Boltzmann constant and T is the absolute temperature). Here, we have also considered that the evolution should produce some entropy.

Finally, considering the initial system state as an equilibrium Gibbs state $\rho_0 = e^{-\beta H}/Z$ (with $Z \equiv \mathrm{Tr} e^{-\beta H}$ and a unitary transformation driving H to H'), we have $p(m|n) = |\langle \psi'_m | U | \psi_n \rangle|^2$. From Eq. (A6), we can obtain $p(m)p(n) = \mathcal{Z} e^{-\beta(E'_m - E_n)} p(m,n)$. Summing it over the initial and final states, it follows that $\sum_{m,n} p(m)p(n) = 1$. So, we can write [30]

$$\begin{aligned}\sum_{m,n} p(m)p(n) &= \sum_{m,n} \mathcal{Z} e^{-\beta(E'_m - E_n)} p(m,n)\\ &= \mathcal{Z} \sum_{m,n} e^{-\beta(E'_m - E_n)} p(n)p(m|n)\\ &= \mathcal{Z} \sum_{m,n} e^{-\beta(E'_m - E_n)} \frac{e^{-\beta E_n}}{Z} |\langle \psi'_m | U | \psi_n \rangle|^2\\ &= \mathcal{Z} \frac{1}{Z} \sum_m e^{-\beta E'_m} \sum_n \langle \psi'_m | U | \psi_n \rangle \langle \psi_n | U^\dagger | \psi'_m \rangle\\ &= \mathcal{Z} \frac{1}{Z} \sum_m e^{-\beta E'_m}\\ &= \mathcal{Z} \frac{Z'}{Z} = 1. \end{aligned} \tag{A10}$$

This implies $\mathcal{Z} = Z/Z'$, where $Z' \equiv \sum_m e^{-\beta E'_m}$. We note that the final state ρ' is not necessarily an equilibrium state in the above development and the system evolution is also not necessarily adiabatic or energy-conserving [30]. Z' works as a partition function for the final Hamiltonian H'. In fact, this quantity will introduce a connection between equilibrium and nonequilibrium system properties.

Defining the Helmholtz free energy as $F \equiv -k_B T \ln Z$, we can write

$$\begin{aligned}\langle \mathcal{I}_{n,m} \rangle &= -\ln \frac{Z}{Z'} + \frac{1}{k_B T} \langle \mathcal{W} \rangle\\ &= \frac{\langle \mathcal{W} \rangle - \Delta F}{k_B T}\\ &= \beta(\langle \mathcal{W} \rangle - \Delta F). \end{aligned} \tag{A11}$$

The preceding equation is compatible with the result obtained by other methods in Ref. [27]. If one takes the averaged exponential of the information density, $\mathcal{I}_{n,m}$, it results in $\langle e^{-\mathcal{I}_{n,m}} \rangle = \langle e^{-\beta(\mathcal{W} - \Delta F)} \rangle = 1$ (according to the previous development), which implies the Jarzinski equality,

$$\langle e^{-\beta \mathcal{W}} \rangle = e^{-\beta \Delta F}. \tag{A12}$$

This result was also pointed out in Ref. [27] using a different approach.

APPENDIX B: DISSIPATION-INFORMATION ACQUISITION INEQUALITY

In this Appendix, we return to our general description of a parameter estimation process described in Fig. 1(a) of the main text. Using the results introduced in the preceding Appendix, we will obtain an inequality for the acquired information in a parameter estimation process and the work dissipated during the process. In this scenario, the mutual information introduced above quantifies the correlations between the probe system in the thermal state (before the initial probe preparation) and the encoded probe (after the encoding process). Let us suppose that during the parameter estimation protocol, the probe system is driven from the thermal ρ_T to an encoded state ρ_φ such that the initial Hamiltonian is $H = \sum_n E_n |\psi_n\rangle\langle\psi_n|$ and the final one reads $H' = \sum_m E'_m |\psi'_m\rangle\langle\psi'_m|$, in a similar way to what was done in the preceding Appendix.

Here, we consider two distinct processes, i.e., the ideal (reversible, theoretical) and the real (irreversible, experimental). The difference between them is that the second one includes a small deviation from the reversible (ideal) dynamics due to fluctuations. We are interested in a high-precision parameter estimation device. Such a device should work closely with reversible dynamics in order to obtain high precision. In that sense, we consider that the actual (real) dynamics includes fluctuations as a slight deviation from ideal dynamics. Below we will consider this scenario.

Defining $p_{0\varphi}(m,n)$ as the joint probability to find the final state in m when the initial state is n, $p_\varphi(m)$ is the final state distribution, and $p_0(n)$ is the initial state distribution, we start by considering the differences in the mutual information densities between the ideal (id) and the real (re) processes, given by

$$\begin{aligned}\mathcal{I}^{\mathrm{re}}_{0:\varphi}(j,k) - \mathcal{I}^{\mathrm{id}}_{0:\varphi}(j,k) &= \ln p^{\mathrm{re}}_{0\varphi}(j,k) - \ln\left[p^{\mathrm{re}}_0(j) p^{\mathrm{re}}_\varphi(k)\right]\\ &\quad - \ln p^{\mathrm{id}}_{0\varphi}(j,k) + \ln\left[p^{\mathrm{id}}_0(j) p^{\mathrm{id}}_\varphi(k)\right]\\ &= \ln \frac{p^{\mathrm{re}}_{0\varphi} t(j,k)}{p^{\mathrm{id}}_{0\varphi}(j,k)} - \ln \frac{p^{\mathrm{re}}_0(j)}{p^{\mathrm{id}}_0(j)}\\ &\quad - \ln \frac{p^{\mathrm{re}}_\varphi(k)}{p^{\mathrm{id}}_\varphi(k)}. \end{aligned} \tag{B1}$$

Multiplying this last equation by $p^{\mathrm{re}}_{0\varphi}(j,k)$ and summing over j and k, we obtain

$$\begin{aligned}&\left\langle \mathcal{I}^{\mathrm{re}}_{0:\varphi}(j,k)\right\rangle_{\mathrm{re}} - \left\langle \mathcal{I}^{\mathrm{id}}_{0:\varphi}(j,k)\right\rangle_{\mathrm{re}}\\ &\quad = \sum_{j,k} p^{\mathrm{re}}_{0\varphi}(j,k)\left[\ln \frac{p^{\mathrm{re}}_{0\varphi}(j,k)}{p^{\mathrm{id}}_{0\varphi}(j,k)} - \ln \frac{p^{\mathrm{re}}_0(j)}{p^{\mathrm{id}}_0(j)} - \ln \frac{p^{\mathrm{re}}_\varphi(k)}{p^{\mathrm{id}}_\varphi(k)}\right]\\ &\quad = S\left(p^{\mathrm{re}}_{0\varphi}||p^{\mathrm{id}}_{0\varphi}\right) - S\left(p^{\mathrm{re}}_0||p^{\mathrm{id}}_0\right) - S\left(p^{\mathrm{re}}_\varphi||p^{\mathrm{id}}_\varphi\right). \end{aligned} \tag{B2}$$

In this last expression, $S(p||q) = \sum_a p(a) \ln[p(a)/q(a)]$ is the relative entropy between distributions p and q. $\langle \cdot \rangle_{\mathrm{re}}$ means that the average is taken over the real (nonideal) process probability distribution.

Using the results of the preceding Appendix, we can write the averaged value of the information density as

$$\left\langle \mathcal{I}^{\mathrm{re}}_{0\varphi}(j,k)\right\rangle_{\mathrm{re}} = \beta(\langle \mathcal{W}^{\mathrm{re}} \rangle - \Delta F^{\mathrm{re}}) = \frac{\langle \mathcal{W}^{\mathrm{re}}_D \rangle}{k_B T}, \tag{B3}$$

with $\langle \mathcal{W}^{\mathrm{re}}_D \rangle$ being the mean dissipated work during the real (nonideal) implementation of the parameter estimation process.

In the next Appendix, we show that

$$S\left(p^{\mathrm{re}}_{0\varphi}||p^{\mathrm{id}}_{0\varphi}\right) \approx \frac{\delta^2_\varphi}{2} \mathcal{F}\left(p^{\mathrm{id}}_{0\varphi}\right), \tag{B4}$$

with δ_φ being the accuracy of the implementation of the process and $\mathcal{F}(p_{0\varphi}^{\rm id})$ the Fisher information of the final ideal distribution (see the main text and the next Appendix for the physical significance and precise mathematical definition of these quantities).

Putting all these results together, we obtain

$$\frac{\langle \mathcal{W}_D^{\rm re}\rangle}{k_B T} - \frac{\delta_\varphi^2}{2}\mathcal{F}\big(p_{0\varphi}^{\rm id}\big) = \Omega,$$

with

$$\Omega = \big\langle \mathcal{I}_{0\varphi}^{\rm id}(j,k)\big\rangle_{\rm re} - S\big(p_0^{\rm re}\big|\big|p_0^{\rm id}\big) - S\big(p_\varphi^{\rm re}\big|\big|p_\varphi^{\rm id}\big). \tag{B5}$$

As a consequence of the inequality in the Cramér-Rao relation (Refs. [20,21]), our main result, Eq. (1) of the main text, is also an inequality. Therefore, to prove it, all we have to do is to prove that $\Omega \geqslant 0$ for all distributions. Thus

$$\begin{aligned}\Omega &= \sum_{j,k} p_{0\varphi}^{\rm re}(j,k)\big[\ln p_{0\varphi}^{\rm id}(j,k) - \ln p_0^{\rm id}(j) - \ln p_\varphi^{\rm id}(k)\big]\\ &\quad - \sum_j p_0^{\rm re}(j)\big[\ln p_0^{\rm re}(j) - \ln p_0^{\rm id}(j)\big]\\ &\quad - \sum_k p_\varphi^{\rm re}(k)\big[\ln p_\varphi^{\rm re}(k) - \ln p_\varphi^{\rm id}(k)\big]\\ &= \sum_{j,k} p_{0\varphi}^{\rm re}(j,k)\ln p_{0\varphi}^{\rm id}(j,k) + H\big(p_0^{\rm re}\big) + H\big(p_\varphi^{\rm re}\big),\end{aligned} \tag{B6}$$

where $H(p) = -\sum_k p(k)\ln p(k)$ is the Shannon entropy. The first term of this expression measures our lack of knowledge about the ideal probability distribution. In fact, $\ln p_{0\varphi}^{\rm id}$ is a measure of the information contained in the ideal distribution (the one we expected to happen). However, events occur according to the real distribution $p_{0\varphi}^{\rm re}$ (due to the finite precision of the experimental apparatus). This is the cause of the loss of information. It is not difficult to show that

$$H\big(p_{0\varphi}^{\rm re}\big) \leqslant -\sum_{j,k} p_{0\varphi}^{\rm re}(j,k)\ln p_{0\varphi}^{\rm id}(j,k) \tag{B7}$$

for every probability distribution. Then, in order to have $\Omega \geqslant 0$, the following relation must be obeyed:

$$\begin{aligned}H\big(p_0^{\rm re}\big) + H\big(p_\varphi^{\rm re}\big) &\geqslant -\sum_{j,k} p_{0\varphi}^{\rm re}(j,k)\ln p_{0\varphi}^{\rm id}(j,k)\\ &\geqslant H\big(p_{0\varphi}^{\rm re}\big),\end{aligned} \tag{B8}$$

which is the well-known superadditivity relation for the entropy. Thus, we are led to conclude that

$$\frac{\langle \mathcal{W}_D^{\rm re}\rangle}{k_B T} - \frac{\delta_\varphi^2}{2}\mathcal{F}\big(p_{0\varphi}^{\rm id}\big) \geqslant 0, \tag{B9}$$

which proves Eq. (9).

APPENDIX C: FISHER INFORMATION AND RELATIVE ENTROPY

One way to define the Fisher information of a probability distribution p_φ is through the calculation of the relative entropy between p_φ and $p_{\varphi+\delta_\varphi}$ yielded by a small shift δ_φ in the parameter φ. δ_φ is then a measure of the accuracy of the process, i.e., the minimum error compatible with the specific process under consideration. It quantifies how much the real probability distribution diverges from the ideal one. We can write the relative entropy $S(p_\varphi\|p_{\varphi+\delta_\varphi})$ as (see Refs. [15,16])

$$\begin{aligned}S(p_\varphi\|p_{\varphi+\delta_\varphi}) &= \sum_{j,k} p_\varphi(j,k)\ln\frac{p_\varphi(j,k)}{p_{\varphi+\delta_\varphi}(j,k)}\\ &= -\sum_{j,k} p_\varphi(j,k)\ln\frac{p_{\varphi+\delta_\varphi}(j,k)}{p_\varphi(j,k)}.\end{aligned} \tag{C1}$$

By using a Taylor expansion about δ_φ, we can write

$$\ln p_{\varphi+\delta_\varphi} - \ln p_\varphi = \delta_\varphi\frac{\partial \ln p_\varphi}{\partial\varphi} + \frac{\delta_\varphi^2}{2}\frac{\partial^2 \ln p_\varphi}{\partial\varphi^2} + O\big(\delta_\varphi^3\big). \tag{C2}$$

Now, substituting (C2) in (C1), the first-order term yields

$$-\delta\sum_{j,k}\frac{p_\varphi(j,k)}{p_\varphi(j,k)}\frac{\partial p_\varphi(j,k)}{\partial\varphi} = -\delta_\varphi\frac{\partial}{\partial\varphi}\sum_{j,k}p_\varphi(j,k) = 0 \tag{C3}$$

while the second-order one leads us to

$$\begin{aligned}&-\frac{\delta_\varphi^2}{2}\sum_{j,k}p_\varphi(j,k)\frac{\partial^2\ln p_\varphi}{\partial\varphi^2}\\ &\quad= -\frac{\delta_\varphi^2}{2}\sum_{j,k}\frac{p_\varphi(j,k)}{p_\varphi(j,k)}\frac{\partial^2 p_\varphi(j,k)}{\partial\varphi^2}\\ &\qquad+\frac{\delta_\varphi^2}{2}\sum_{j,k}p_\varphi(j,k)\left(\frac{1}{p_\varphi(j,k)}\frac{\partial p_\varphi(j,k)}{\partial\varphi}\right)^2.\end{aligned} \tag{C4}$$

Since the probability distribution is normalized,

$$\sum_{j,k}\frac{\partial^2 p_\varphi}{\partial\varphi^2} = 0, \tag{C5}$$

and keeping terms up to second order, we obtain

$$\begin{aligned}S\big(p_\varphi\big|\big|p_{\varphi+\delta_\varphi}\big) &\approx \frac{\delta_\varphi^2}{2}\sum_{j,k}p_\varphi(j,k)\left(\frac{\partial}{\partial\varphi}\ln p_\varphi(j,k)\right)^2\\ &= \frac{\delta_\varphi^2}{2}\mathcal{F}(p_\varphi),\end{aligned} \tag{C6}$$

where $\mathcal{F}(p_\varphi)$ is the Fisher information of p_φ. This approximation implies that we must have

$$\left|\frac{\delta_\varphi}{\varphi}\right| \ll 1, \tag{C7}$$

meaning that the error in the measurement is much smaller than the parameter we are measuring. This is quite a reasonable assumption since an error of the same order of magnitude of the parameter would render the measurement process meaningless.

To prove Eq. (B4), let us introduce the *Jeffreys divergence* (Ref. [15]),

$$J(p_\varphi, p_{\varphi+\delta_\varphi}) = S(p_\varphi || p_{\varphi+\delta_\varphi}) + S(p_{\varphi+\delta_\varphi} || p_\varphi) = \sum_{j,k} [p_\varphi(j,k) - p_{\varphi+\delta_\varphi}(j,k)] \ln \frac{p_\varphi(j,k)}{p_{\varphi+\delta_\varphi}(j,k)}. \tag{C8}$$

Since (C8) is symmetrical, we may rewrite it as

$$J(p_\varphi, p_{\varphi+\delta_\varphi}) = \sum_{j,k} \Delta p_{\varphi+\delta_\varphi}(j,k) \ln \left(1 + \frac{\Delta p_{\varphi+\delta_\varphi}(j,k)}{p_\varphi(j,k)}\right), \tag{C9}$$

where $\Delta p_{\varphi+\delta_\varphi} = p_{\varphi+\delta_\varphi} - p_\varphi$. Due to the fact that δ_φ is small, the distribution $p_{\varphi+\delta_\varphi}$ will be close to p_φ. Therefore,

$$\ln \left(1 + \frac{\Delta p_{\varphi+\delta_\varphi}(j,k)}{p_\varphi(j,k)}\right) \approx \frac{\Delta p_{\varphi+\delta_\varphi}(j,k)}{p_\varphi(j,k)}, \tag{C10}$$

allowing us to write

$$\begin{aligned} J(p_\varphi, p_{\varphi+\delta_\varphi}) &\approx \sum_{j,k} p_\varphi(j,k) \left(\frac{\Delta p_{\varphi+\delta_\varphi}(j,k)}{p_\varphi(j,k)}\right)^2 \\ &= \sum_{j,k} p_\varphi(j,k) \left(\frac{\delta_\varphi}{p_\varphi} \frac{\Delta p_{\varphi+\delta_\varphi}(j,k)}{\delta_\varphi}\right)^2 \\ &= \delta_\varphi^2 \sum_{j,k} p_\varphi(j,k) \left(\frac{1}{p_\varphi(j,k)} \frac{\partial p_\varphi(j,k)}{\partial \varphi}\right)^2 \\ &= \delta_\varphi^2 \mathcal{F}(p_\varphi). \end{aligned} \tag{C11}$$

From Eq. (C8) we have $S(p_{\varphi+\delta_\varphi} \| p_\varphi) = J(p_\varphi, p_{\varphi+\delta_\varphi}) - S(p_\varphi \| p_{\varphi+\delta_\varphi})$, so

$$S(p_{\varphi+\delta_\varphi} \| p_\varphi) \approx \frac{\delta_\varphi^2}{2} \mathcal{F}(p_\varphi). \tag{C12}$$

Finally, making the identifications $p_\varphi \equiv p_{0\varphi}^{\rm id}$ and $p_{\varphi+\delta_\varphi} \equiv p_{0\varphi}^{\rm re}$, we obtain Eq. (B4), thus proving our main claim, Eq. (1) of the main text.

[1] J. C. Maxwell, *Theory of Heat* (Longmans, Green and Co., London, 1871).
[2] L. Brillouin, J. Appl. Phys. **22**, 334 (1951).
[3] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[4] L. del Rio, J. Åberg, R. Renner, O. C. O. Dahlsten, and V. Vedral, Nature (London) **474**, 61 (2011).
[5] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Nature (London) **483**, 187 (2012).
[6] T. Sagawa and M. Ueda, Phys. Rev. Lett. **100**, 080403 (2008).
[7] T. Sagawa and M. Ueda, Phys. Rev. Lett. **102**, 250602 (2009).
[8] T. Sagawa and M. Ueda, Phys. Rev. Lett. **104**, 090602 (2010).
[9] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Nat. Phys. **6**, 988 (2010).
[10] T. Sagawa and M. Ueda, in *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*, edited by R. Klages, W. Just, and C. Jarzynski (Wiley-VCH, Weinheim, 2012).
[11] S. Hormoz, Phys. Rev. E **87**, 022129 (2013).
[12] E. Hänggi and S. Wehner, Nat. Commun. **4**, 1670 (2013).
[13] R. A. Fisher, Philos. Trans. R. Soc. London, Ser. A **222**, 309 (1922).
[14] R. A. Fisher, Proc. Cambridge Philos. Soc. **22**, 700 (1925).
[15] S. Kullback, *Information Theory and Statistics* (Dover, New York, 1967).
[16] M. Hayashi, *Quantum Information* (Springer-Verlag, Berlin, 2006).
[17] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[18] S. Popescu, A. J. Short, and A. Winter, Nat. Phys. **2**, 754 (2006).
[19] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, UK, 1980).
[20] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).
[21] C. R. Rao, *Linear Statistical Inference and Its Application* (Wiley, New York, 1973).
[22] J. Oppenheim and S. Wehner, Science **330**, 1072 (2010).
[23] O. C. O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, New. J. Phys. **13**, 053015 (2011).
[24] M. Planck, Ann. Phys. **306**, 69 (1900).
[25] G. E. A. Matsas, V. Pleitez, A. Saa, and D. A. T. Vanzella, arXiv:0711.4276.
[26] A. Bohr, B. R. Mottelson, and O. Ulfbeck, Proc. Natl. Acad. Sci. (U.S.A.) **105**, 17301 (2008).
[27] V. Vedral, J. Phys. A **45**, 272001 (2012).
[28] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102 (2007).
[29] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, New York, 2003).
[30] H. Tasaki, arXiv:cond-mat/0009244.