Open access • Journal Article • DOI:10.1103/PHYSREVA.86.044302

Holevo's bound from a general quantum fluctuation theorem — Source link

Dvir Kafri, Sebastian Deffner

Institutions: University of Maryland, College Park

Published on: 22 Oct 2012 - Physical Review A (American Physical Society (APS))

Topics: Holevo's theorem, Classical capacity, Fluctuation theorem, Quantum fluctuation and Quantum statistical mechanics

Holevo's bound from a general quantum fluctuation theorem

Dvir Kafri[1] and Sebastian Deffner[2]

[1]*Joint Quantum Institute, Department of Physics, University of Maryland, College Park, Maryland 20742, USA*

[2]*Department of Chemistry and Biochemistry and Institute for Physical Sciences and Technology, University of Maryland, College Park, Maryland 20742, USA*

(Received 23 August 2012; revised manuscript received 5 October 2012; published 22 October 2012)

We give a derivation of Holevo's bound using an important result from nonequilibrium statistical physics, the fluctuation theorem. To do so we develop a general formalism of quantum fluctuation theorems for two-time measurements, which explicitly accounts for the back action of quantum measurements as well as possibly nonunitary time evolution. For a specific choice of observables this fluctuation theorem yields a measurement-dependent correction to the Holevo bound, leading to a tighter inequality. We conclude by analyzing equality conditions for the improved bound.

DOI: 10.1103/PhysRevA.86.044302 PACS number(s): 03.67.Hk, 05.30.−d, 05.40.−a, 89.70.Kn

Thermodynamics and information theory are intimately connected. The most prominent evidence for this relationship is that the Clausius entropy [1] is given by the Shannon information [2] in systems at thermal equilibrium [3]. In particular, Landauer's principle [4,5] illustrates that information is a physical, measurable quantity. Thermodynamic work has to be performed in order to create or erase information. Landauer's principle can therefore be understood as a statement of the second law of thermodynamics in an information theoretic context. This is also true for Holevo's bound [6], which limits the amount of classical information that can be encoded in a generic quantum system. Recently, nanodevices, for which these principles are directly applicable, have become experimentally accessible [7]. These controlled quantum systems have applications ranging from quantum simulation [8,9], cryptography [10], and computing [11–13] to metrology [14–17]. The main obstacles towards realization of such devices are control noise and interactions with the environment. Thus, a thermodynamic study is necessary to fully understand their information theoretic properties. However, most nanodevices operate far from thermal equilibrium, so tools from nonequilibrium statistical physics are required. In recent years, formulations of the second law have been derived which are valid arbitrarily far from equilibrium. These so-called fluctuation theorems, in particular, the Jarzynski equality [18], enable the calculation of equilibrium quantities from nonequilibrium averages over many realizations of a single process. They also encompass nonequilibrium, information-theoretic generalizations of the second law. As Landauer's principle is a direct implication of this approach [19–21], one may ask whether Holevo's bound is also such a result. One complication in this case is that the approach to fluctuation theorems for quantum systems is mathematically and conceptually more involved. Thermodynamic quantities, which are not given as state functions, cannot be assigned a Hermitian operator [22]. The proper formulation of quantum thermodynamics for nonequilibrium systems, especially quantum fluctuation theorems [23], must therefore be treated with care.

The purpose of the present Brief Report is twofold. In the first part we derive a general quantum fluctuation theorem that accounts for the back action of measurements on reduced systems (see Ref. [24] for a similar approach). To this end, we consider an experimental point of view; we assume that the system of interest is coupled to an environment which is experimentally inaccessible. Such measurements on open quantum systems are inherently incomplete since they ignore environmental degrees of freedom. Information is *lost* that, in principle, could have been acquired by concurrent measurement of the reservoir. A general formulation of quantum fluctuation theorems must explicitly account for these effects. The integral fluctuation theorem we derive is applicable to arbitrary orthogonal measurements for systems undergoing both unitary and nonunitary dynamics.

In the second part we focus on an information-theoretic consequence of the general quantum fluctuation theorem: Holevo's theorem. The derivation and implications of this result have attracted much attention [25–29]. Like more recent works [30–32], our derivation results in a sharpened statement of Holevo's bound, which takes into account the choice of measurement used to obtain the encoded information. Further our treatment is based on results weaker than the monotonicity of relative entropy and directly leads to necessary and sufficient equality conditions. This illustrates an interesting connection between quantum thermodynamics and quantum information theory.

General quantum fluctuation theorem. Consider a time-dependent quantum system $\mathcal{S}$ with Hilbert space $\mathcal{H}_\mathcal{S}$ and initial density matrix ρ_0. Information about the state of the system is obtained by performing measurements on $\mathcal{S}$ at the beginning and end of a specific process. Initially, a quantum measurement is made of observable A^{i}, with eigenvalues a^{i}_m. Letting Π^{i}_m denote the orthogonal projectors into the eigenspaces of A^{i}, we have $A^{\mathrm{i}} = \sum_m a^{\mathrm{i}}_m \Pi^{\mathrm{i}}_m$. Note that the eigenvalues $\{a^{\mathrm{i}}_m\}$ can be degenerate, so the projectors $\{\Pi^{\mathrm{i}}_m\}$ may have a rank greater than 1. Unlike the classical case, as long as ρ_0 and A^{i} do not have a common set of eigenvectors, i.e., they do not commute, performing a measurement on $\mathcal{S}$ alters its statistics. Measuring a^{i}_m maps ρ_0 to the state $\Pi^{\mathrm{i}}_m \rho_0 \Pi^{\mathrm{i}}_m / p_m$, where $p_m = \mathrm{tr}\{\Pi^{\mathrm{i}}_m \rho_0 \Pi^{\mathrm{i}}_m\}$ is the probability of the measurement outcome a^{i}_m. Generally accounting for all possible measurement outcomes, the statistics of $\mathcal{S}$ after the measurement are given by the weighted average of all projections,

$$M^{\mathrm{i}}(\rho_0) = \sum_m \Pi^{\mathrm{i}}_m \, \rho_0 \, \Pi^{\mathrm{i}}_m \, . \tag{1}$$

If ρ_0 commutes with $A^{\rm i}$, it commutes with each $\Pi^{\rm i}_m$, so $M^{\rm i}(\rho_0) = \sum_m \Pi^{\rm i}_m \Pi^{\rm i}_m \rho_0 = \rho_0$, and the statistics of the system are unaltered by the measurement. After measuring $a^{\rm i}_m$, $\mathcal{S}$ undergoes a generic time evolution, after which it is given by $\mathbb{E}(\Pi^{\rm i}_m \rho_0 \Pi^{\rm i}_m)/p_m$. Here $\mathbb{E}$ represents any linear (unitary or nonunitary) quantum transformation, which is trace preserving and maps non-negative operators to non-negative operators. Further, we require that this holds whenever $\mathbb{E}$ is extended to an operation $\mathbb{E} \otimes \mathbb{I}_{\mathcal{E}}$ on any enlarged Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_{\mathcal{E}}$ ($\mathbb{I}_{\mathcal{E}}$ being the identity map on $\mathcal{H}_{\mathcal{E}}$). Such a transformation is called a trace-preserving, completely positive (TCP) map [33]. After this evolution, a measurement of a second (not necessarily the same) observable, $A^{\rm f} = \sum_n a^{\rm f}_n \Pi^{\rm f}_n$, is performed on $\mathcal{S}$. The probability of measuring $a^{\rm f}_n$, conditioned on having first measured $a^{\rm i}_m$, is $p_{n|m} = {\rm tr}\{\Pi^{\rm f}_n \mathbb{E}(\Pi^{\rm i}_m \rho_0 \Pi^{\rm i}_m)\}/p_m$. Accordingly, the joint probability distribution $p_{m\to n}$ reads

$$p_{m\to n} = p_m p_{n|m} = {\rm tr}\left\{\Pi^{\rm f}_n \, \mathbb{E}\left(\Pi^{\rm i}_m \rho_0 \Pi^{\rm i}_m\right)\right\}. \tag{2}$$

We are interested in the probability distribution of possible measurement outcomes, $\mathcal{P}(\Delta a) = \langle \delta(\Delta a - \Delta a_{n,m})\rangle$, where $\Delta a_{n,m} = a^{\rm f}_n - a^{\rm i}_m$ is a random variable determined in a single measurement run. Its probability distribution is given by averaging over all possible realizations,

$$\mathcal{P}(\Delta a) = \sum_{m,n} \delta(\Delta a - \Delta a_{n,m})\, p_{m\to n}\,. \tag{3}$$

To derive the integral fluctuation theorem we follow the standard approach and compute its characteristic function, $\mathcal{G}(s)$, which is the Fourier transform of $\mathcal{P}(\Delta a)$ [23]:

$$\begin{aligned}\mathcal{G}(s) &= \int d(\Delta a)\, \mathcal{P}(\Delta a)\, \exp(is\,\Delta a) \\ &= {\rm tr}\{\exp(isA^{\rm f})\, \mathbb{E}(M^{\rm i}(\rho_0) \exp(-isA^{\rm i}))\}\,. \end{aligned}\tag{4}$$

Choosing $s = i$, we obtain the general quantum fluctuation theorem

$$\langle \exp(-\Delta a)\rangle = \gamma\,. \tag{5}$$

Since it is explicitly dependent on the map $\mathbb{E}$, the quantity γ accounts for the information *lost* by not measuring the environment. It plays a crucial role in the following discussion and is given by

$$\gamma = {\rm tr}\{\exp(-A^{\rm f})\, \mathbb{E}(M^{\rm i}(\rho_0) \exp(A^{\rm i}))\}\,. \tag{6}$$

Similar fluctuation theorems of the form $\langle \exp(-\Sigma)\rangle = \gamma_{\rm cl}$ have been derived in the context of classical feedback processes, where Σ is an entropy production [34]. We note that by appropriate choice of initial and final observables $A^{\rm i}$ and $A^{\rm f}$, Eq. (6) reproduces many known quantum fluctuation theorems [22–24,35–39], which we will discuss in detail elsewhere.

A complementary result to the fluctuation theorem is Jensen's inequality, which states that for any convex function $\phi''(x) \geqslant 0$ and random variable x, $\langle \phi(x)\rangle \geqslant \phi(\langle x\rangle)$ [40]. Applying this to Eq. (5) yields

$$\langle \Delta a\rangle \geqslant -\ln(\gamma)\,. \tag{7}$$

For specific choices of thermodynamically relevant observables $A^{\rm i}$ and $A^{\rm f}$, this relation can be understood as a formulation of the Clausius inequality. In particular, for a unitary time evolution $U_\tau = \mathcal{T}_> \exp[-i\int_0^\tau H(t)dt]$, an initial Gibbsian state $\rho_0 = \exp[-\beta H(0)]/Z_0$, and corresponding energy measurements, $A^{\rm i} = \beta H(0)$ and $A^{\rm f} = \beta H(\tau)$, Eq. (5) reproduces the quantum Jarzynski equality [22,41,42]. Accordingly, Eq. (7) reduces to the maximum work theorem, $\beta\langle W\rangle \geqslant \beta\Delta F$, where $\langle W\rangle = \langle H(\tau)\rangle - \langle H(0)\rangle = \langle \Delta a\rangle/\beta$ and $\beta\Delta F = -\ln(Z_\tau/Z_0) = -\ln(\gamma)$.

Holevo's bound. We now use the fluctuation theorem (5) to derive a sharpened version of Holevo's bound. This bound sets a limit on how much classical information can be sent through a (noisy) quantum channel. Let us consider a message composed of code words w_j that appear with probability π_j. A messenger (Alice) attempts to transfer this message to a receiver (Bob) by encoding each word w_j in a quantum state and transmitting that state to Bob. We assume that Bob receives the state ρ_j, which may have come through a lossy medium and therefore may be different from the original state prepared by Alice. Bob attempts to infer the word w_j from the encoding by making a generalized measurement of the state ρ_j. This corresponds to introducing a probe, initially in a pure state $|0\rangle$, and making an orthogonal measurement on the compound state $\rho_j \otimes |0\rangle\langle 0|$ [43]. If $\{\Pi_k\}$ represents the set of orthogonal projectors corresponding to Bob's measurement, the probability of measuring Π_k, given message w_j, is given by

$$\pi_{k|j} = {\rm tr}\{\rho_j \otimes |0\rangle\langle 0|\Pi_k\} = {\rm tr}\{\rho_j M_k\}\,, \tag{8}$$

where $M_k = \langle 0|\Pi_k|0\rangle$ are operators acting only on the encoding degree of freedom. Although the operators $\{M_k\}$ are non-negative and $\sum_k M_k = \mathbb{I}$, they are generally not projectors, $M_k^2 \neq M_k$. Such a collection $\{M_k\}$ is called a positive operator-valued measure (POVM) and describes the most general measurement on a quantum system. The classical message distribution $\{\pi_j\}$, output quantum encoding $\{\rho_j\}$, and POVM elements $\{M_k\}$ define a classical-quantum channel [44].

A proper measure of how well Bob decodes Alice's message is the mutual information between the encoded message and measurement distributions, $I = \sum_{jk} \pi_j\, \pi_{k|j}\, \ln(\pi_{k|j}/\pi_k)$, where $\pi_k = \sum_j \pi_j\, \pi_{k|j}$ is the overall probability of measuring Π_k. Note that $I = \sum_j \pi_j D(\pi_{k|j}||\pi_k)$, where $D(\pi_{k|j}||\pi_k) = \sum_k \pi_{k|j} \ln(\pi_{k|j}/\pi_k)$ is the (classical) relative entropy [45]. Hence I is a sum of non-negative terms and is zero if and only if $\pi_{k|j} = \pi_k$ for all k, j. That is, I vanishes only if all outcomes of the measurement are independent of the encoded word, so that Bob always learns nothing about the message.

The probability of the message being w_j, conditioned on Bob measuring Π_k, is $\pi_{j|k} = \pi_{k|j}\, \pi_j/\pi_k$. We have

$$I = S(\{\pi_j\}) + \sum_k \pi_k \sum_j \pi_{j|k} \ln \pi_{j|k}, \tag{9}$$

where $S(\{\pi_j\}) = -\sum_j \pi_j \ln \pi_j$ is the Shannon information of distribution $\{\pi_j\}$. Since $x\ln(x) \leqslant 0$ for $|x| \leqslant 1$, with equality only for $x = 0$ and $x = 1$, we observe that $I \leqslant S(\{\pi_j\})$, with equality if and only if $\pi_{j|k}$ is 0 or 1 for all j,k. In other words, the mutual information I is at most $S(\{\pi_j\})$, with equality if and only if Bob correctly decodes the message

in every instance. If Alice's encoded states are not perfectly distinguishable (that is, if the supports of ρ_j and $\rho_{j'}$ are not orthogonal for some $j \neq j'$), then I can never equal $S(\{\pi_j\})$, no matter what measurement Bob chooses to make. Holevo's theorem is then an upper bound for I, namely,

$$\chi \equiv S(\bar{\rho}) - \sum_j \pi_j S(\rho_j) \geqslant I, \tag{10}$$

where $S(\rho) = -\mathrm{tr}\{\rho \ln \rho\}$ is the von Neumann entropy and $\bar{\rho} = \sum_j \pi_j \rho_j$ is the density matrix describing the statistics of the encoding given no knowledge of the message word. Heuristically, the Holevo quantity χ can be considered as the uncertainty of the encoding with no knowledge of the message minus the average remaining uncertainty given knowledge of the message. Note that if the encoded states are distinguishable, i.e., $\rho_j \rho_{j'} = \mathbb{O}$ for all $j \neq j'$, then $\chi = S(\{\pi_j\})$, so that with a proper measurement Bob may always correctly decode Alice's message.

We now show how Holevo's theorem (10) follows as a consequence of the general quantum fluctuation theorem (5). To do this we must appropriately choose the initial state, evolution operation, and observables so that the random variable $\langle \Delta a \rangle$ averages to $\chi - I$. In the language of the general fluctuation theorem, let the initial state ρ_0 reside in a composite Hilbert space $\mathcal{H}_\mathcal{E} \otimes \mathcal{H}_\mathcal{P} \otimes \mathcal{H}_\mathcal{M}$. $\mathcal{H}_\mathcal{E}$ represents the encoding Hilbert space, which Alice prepares and Bob then measures, $\mathcal{H}_\mathcal{P}$ is the probe Hilbert space accessible only to Bob, and $\mathcal{H}_\mathcal{M}$ is a *message* Hilbert space. Note that $\mathcal{H}_\mathcal{M}$ is not a *real*, physically accessible subspace, but rather a mathematical construction denoting the *memory* for the classical information of the message [44]. We have

$$\rho_0 = \sum_j \pi_j \rho_j \otimes |0\rangle\langle 0| \otimes |j\rangle\langle j|, \tag{11}$$

where the states $|j\rangle$, each corresponding to word w_j, form an orthonormal basis for $\mathcal{H}_\mathcal{M}$. We see that, with probability π_j, ρ_0 corresponds to the *message state* $|j\rangle$. The measured initial and final observables are

$$\begin{aligned} A^\mathrm{i} &= \sum_j \ln\left(\hat{\rho}_j^{-1}\right) \otimes |0\rangle\langle 0| \otimes |j\rangle\langle j|, \\ A^\mathrm{f} &= -\ln(\bar{\rho} \otimes |0\rangle\langle 0|) \otimes \mathbb{I}_\mathcal{M} - \sum_{k,j} I_{k,j}\, \Pi_k \otimes |j\rangle\langle j|, \end{aligned} \tag{12}$$

where $I_{k,j} = \ln(\pi_{k|j}/\pi_k)$. Here $\hat{\rho}_j^{-1}$ denotes the inverse within the support of ρ_j, so that $\hat{\rho}_j^{-1}|\psi\rangle = 0$ whenever $\rho_j|\psi\rangle = 0$. This form for Eq. (12) ensures that $\exp(A^\mathrm{i})$ and $\exp(-A^\mathrm{f})$ are bounded operators.

Note that the states ρ_j do not represent the original encoding set up by Alice, but rather its time-evolved state after undergoing dynamics in a quantum channel. To apply the fluctuation theorem (5), we start with the output of this channel and perform the two measurements, A^i and A^f, immediately after each other. The TCP map crucial for Eq. (5) is thus the identity map $\mathbb{E}(\rho) = \rho$. A^i commutes with ρ_0, so after measurement of A^i, measurement of A^f is carried out on the *same* state, $M^\mathrm{i}(\rho_0) = \rho_0$. Computing $\langle \Delta a \rangle = \mathrm{tr}\{(A^\mathrm{f} - A^\mathrm{i})\rho_0\} = \chi - I$, Eq. (7) is

$$\chi - I \geqslant -\ln(\gamma), \tag{13}$$

where the corresponding quantum efficacy is given by

$$\gamma = \mathrm{tr}\{\exp(-A^\mathrm{f})\rho_0 \exp(A^\mathrm{i})\}. \tag{14}$$

Equations (13) and (14) constitute the sharpened Holevo's bound as a consequence of the general quantum fluctuation theorem (5). Indeed, our new bound is tighter than the usual inequality (10), in the sense that the correction term, $-\ln(\gamma)$, is always non-negative. Consider

$$\begin{aligned} \rho_0 \exp(A^\mathrm{i}) &= \sum_j \pi_j \rho_j \exp\left[\ln\left(\hat{\rho}_j^{-1}\right)\right] \otimes |0\rangle\langle 0| \otimes |j\rangle\langle j| \\ &= \sum_j \pi_j \hat{P}_j \otimes |0\rangle\langle 0| \otimes |j\rangle\langle j|, \end{aligned} \tag{15}$$

where $\hat{P}_j$ is the projector into the support of ρ_j. We can rewrite Eq. (14) with Eq. (15) as

$$\begin{aligned} \gamma &= \mathrm{tr}\{\exp(-A^\mathrm{f})\, \rho_0 \exp(A^\mathrm{i})\} \\ &= \mathrm{tr}\left\{\exp(-A^\mathrm{f}) \sum_j \pi_j \hat{P}_j \otimes |0\rangle\langle 0| \otimes |j\rangle\langle j|\right\} \\ &\leqslant \sum_j \pi_j\, \mathrm{tr}\left\{\exp\left(\ln(\bar{\rho} \otimes |0\rangle\langle 0|) + \sum_k I_{k,j}\, \Pi_k\right)\right\}, \end{aligned} \tag{16}$$

where the inequality is justified by noting that $\exp[\ln(\bar{\rho} \otimes |0\rangle\langle 0|) + \sum_k I_{k,j}\, \Pi_k]$ is non-negative and $\hat{P}_j \otimes |0\rangle\langle 0|$ is a projection operator. We now use a statement of the Golden-Thompson inequality [46,47]; that is, for any Hermitian operators A and B, we have $\mathrm{tr}\{\exp(A+B)\} \leqslant \mathrm{tr}\{\exp(A)\exp(B)\}$. Note that in the present case, A and B are both logarithms of bounded Hermitian operators and are only bounded from *above*, although the Golden-Thompson inequality still holds [48]. Accordingly, we have

$$\begin{aligned} \gamma &= \mathrm{tr}\{\exp(-A^\mathrm{f})\, \rho_0 \exp(A^\mathrm{i})\} \\ &\leqslant \sum_j \pi_j\, \mathrm{tr}\left\{\exp[\ln(\bar{\rho} \otimes |0\rangle\langle 0|)] \exp\left(\sum_k I_{k,j}\, \Pi_k\right)\right\} \\ &= \sum_j \pi_j\, \mathrm{tr}\left\{(\bar{\rho} \otimes |0\rangle\langle 0|) \sum_k \pi_{k|j}/\pi_k\, \Pi_k\right\}. \end{aligned} \tag{17}$$

From the definition $\pi_k = \sum_j \pi_j \pi_{k|j}$ we finally obtain

$$\gamma \leqslant \mathrm{tr}\left\{\bar{\rho} \otimes |0\rangle\langle 0| \sum_k \Pi_k\right\} = 1, \tag{18}$$

which shows that $-\ln(\gamma) \geqslant 0$, as desired. We note that our derivation does not invoke the monotonicity of the relative

entropy or equivalent statements [49]. Instead, we have used only Jensen's inequality and the Golden-Thompson inequality, which are weaker results [50,51].

Equality conditions. Holevo's bound (12) is obtained with the help of Jensen's inequality. For strictly convex functions $\phi''(x) > 0$, the Jensen bound $\langle\phi(x)\rangle \geqslant \phi(\langle x\rangle)$ achieves equality if and only if the random variable x is constant valued. This allows us to derive the equality conditions for (13) in a straightforward manner. Specifically, equality is achieved only if

$$\left(-\ln\left(\hat{\rho}_j^{-1}\right) - \ln(\bar{\rho}) - \sum_k I_{k,j} M_k\right)\hat{P}_j = -\ln(\gamma)\hat{P}_j \qquad (19)$$

for all j. This follows from a few simple observations.

First, assume that $\chi - I = -\ln(\gamma)$. As A^{i}, ρ_0, and the projectors $\mathbb{I}\otimes|0\rangle\langle 0|\otimes|j\rangle\langle j|$ mutually commute, we consider a mutual eigenprojector $R_{mj} = R_m^{(j)}\otimes|0\rangle\langle 0|\otimes|j\rangle\langle j|$ such that $A^{\mathrm{i}}R_{mj} = a_m^{\mathrm{i}}R_{mj}$ and $\rho_0 R_{mj} \neq 0$. Since the function $\exp(x)$ is strictly convex, the random variable Δa_{nm} obtained from the measurements of A^{i} and A^{f} has to satisfy $\Delta a_{nm} = -\ln(\gamma)$ for all measurements with nonzero probability. Hence an initial measurement of a_m^{i} implies with certainty a final measurement $a_m^{\mathrm{i}} - \ln(\gamma)$. Since R_{mj} is a projector into an eigenspace of ρ_0, any state satisfying $R_{mj}|\psi\rangle = |\psi\rangle$ must therefore also be an eigenstate of A^{f} with eigenvalue $a_m^{\mathrm{i}} - \ln(\gamma)$, so

$$(A^{\mathrm{f}} - A^{\mathrm{i}})R_{mj} = -\ln(\gamma)R_{mj}. \qquad (20)$$

Using definition (12) and $M_k = \langle 0|\Pi_k|0\rangle$, Eq. (19) follows by summing on m, noting that $\sum_m R_{mj} = \hat{P}_j\otimes|0\rangle\langle 0|\otimes|j\rangle\langle j|$.

Conversely, assume that Eq. (19) holds for all j. Since A^{i} and ρ_0 commute, we have

$$\begin{aligned}\chi - I &= \mathrm{tr}\{\rho_0(A^{\mathrm{f}} - A^{\mathrm{i}})\}\\ &= \sum_j \pi_j \mathrm{tr}\left\{\rho_j\left(\ln(\rho_j) - \ln(\bar{\rho}) - \sum_k I_{k,j}M_k\right)\hat{P}_j\right\}\\ &= \sum_j \pi_j \mathrm{tr}\{\rho_j[-\ln(\gamma)\hat{P}_j]\} = -\ln(\gamma)\sum_j \pi_j\\ &= -\ln(\gamma) \qquad (21)\end{aligned}$$

We conclude that Eq. (19) is equivalent to equality in Eq. (13). Observe that since $\chi - I \geqslant -\ln(\gamma) \geqslant 0$, the equality condition for $\chi = I$, Eq. (19) with $\ln(\gamma) = 0$, is obtained as a corollary of our result [49]. Equation (19) may be used to determine the bound saturating observable A^{f} self-consistently.

Concluding remarks. We developed a general framework for quantum fluctuation theorems by explicitly accounting for the back action of quantum measurements. With this result, we showed that quantum-mechanical formulations of the second law are intimately tied to quantum information theory by deriving Holevo's bound as a consequence of a fluctuation theorem. The new approach provides not only simple derivation but also a sharpened statement of the original bound and a corresponding equality criterion.

Acknowledgments. The authors thank Jacob Taylor and Eric Lutz for interesting discussions. SD acknowledges financial support by a fellowship within the postdoc-program of the German Academic Exchange Service (DAAD, contract No D/11/40955). DK acknowledges financial support by a fellowship from the Joint Quantum Institute.

[1] R. Clausius, *Abhandlungen über die mechanische Wärmetheorie* (Vieweg, Braunschweig, Germany, 1864).
[2] C. E. Shannon, Bell Syst. Technol. J. **27**, 379 (1948).
[3] H. Callen, *Thermodynamics and an Introduction to Thermostastistics* (Wiley, New York, 1985).
[4] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[5] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenscheinder, and E. Lutz, Nature (London) **483**, 187 (2012).
[6] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[7] C. Monroe, Nature (London) **416**, 238 (2002).
[8] R. Feynman, Int. J. Theor. Phys. **21**, 476 (1982).
[9] S. Lloyd, Science **273**, 1073 (1996).
[10] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[11] L. K. Grover Phys. Rev. Lett. **79**, 325 (1997).
[12] P. W. Shor, SIAM Rev. **41**, 303 (1999).
[13] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).
[14] M. Kasevich and S. Chu, Phys. Rev. Lett. **67**, 181 (1991).
[15] H. J. Kimble, Y. Levin, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin, Phys. Rev. D **65**, 022002 (2001).
[16] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[17] J. Hudson, D. M. Kara, I. J. Smallman, B. E. Sauer, M. R. Tarbutt, and E. A. Hinds, Nature (London) **473**, 493 (2011).
[18] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[19] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007).
[20] C. Jarzynski, Eur. Phys. J. B **64**, 331 (2008).
[21] M. Esposito and C. van den Broeck, Europhys. Lett. **95**, 40004 (2011).
[22] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102(R) (2007).
[23] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).
[24] V. Vedral, J. Phys. A **45**, 272001 (2012).
[25] H. P. Yuen and M. Ozawa, Phys. Rev. Lett. **70**, 363 (1993).
[26] M. J. W. Hall and M. J. O'Rourke, Quantum Opt. **5**, 161 (1993).
[27] C. A. Fuchs and C. M. Caves, Phys. Rev. Lett. **73**, 3047 (1994).
[28] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[29] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[30] B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. Lett. **76**, 3452 (1996).

[31] K. Jacobs, Phys. Rev. A **68**, 054302 (2003).
[32] K. Jacobs, J. Math. Phys. **47**, 012102 (2006).
[33] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).
[34] For these the right-hand side, $\gamma_{\rm cl}$, is commonly called the (classical) efficacy of the feedback protocol [39,52]. In the present context γ can then be interpreted as a quantum efficacy corresponding to the observables $A^{\rm i}$ and $A^{\rm f}$ and the TCP map $\mathbb{E}$. Note that γ only takes the form of the classical efficacy and usually may not be regarded as its physical quantum analog [53].
[35] P. Talkner, M. Campisi, and P. Hänggi, J. Stat. Mech. (2009) P02025.
[36] S. Deffner and E. Lutz, Phys. Rev. Lett. **107**, 140404 (2011).
[37] M. Campisi, P. Talkner, and P. Hänggi, Phys. Rev. E **83**, 041114 (2011).
[38] C. Jarzynski and D. K. Wójcik, Phys. Rev. Lett. **92**, 230602 (2004).
[39] Y. Morikuni and H. Tasaki, J. Stat. Phys. **143**, 1 (2011).
[40] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, Oxford, 1987).
[41] J. Kurchan, arXiv:cond-mat/0007360 (2000).
[42] H. Tasaki, arXiv:cond-mat/0009244 (2000).
[43] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, 1992).
[44] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[45] S. Kullback, *Information Theory and Statistics* (Peter Smith, Gloucester, MA, 1978).
[46] S. Golden, Phys. Rev. **137**, B1127 (1965).
[47] C. J. Thompson, J. Math. Phys. **6**, 1812 (1965).
[48] M. B. Ruskai, Commun. Math. Phys. **26**, 280 (1972).
[49] M. B. Ruskai, J. Math. Phys. **9**, 4358 (2002).
[50] D. Petz, Commun. Math. Phys. **114**, 345 (1988).
[51] D. Petz, in *Functional Analysis and Operator Theory*, Banach Center Publications, Vol. 30 (Institute of Mathematics, Polish Academy of Sciences, Warsaw, 1994).
[52] T. Sagawa and M. Ueda, Phys. Rev. Lett. **104**, 090602 (2010).
[53] T. Sagawa and M. Ueda, Phys. Rev. E **85**, 021104 (2012).