

Limitation on the Amount of Accessible Information in a Quantum Channel

Benjamin Schumacher,^{1,2} Michael Westmoreland,³ and William K. Wootters⁴

¹Department of Physics, Kenyon College, Gambier, Ohio 43022

²Theoretical Astrophysics, T-6, MS B288, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

³Department of Mathematics, Denison University, Granville, Ohio 43023

⁴Department of Physics, Williams College, Williamstown, Massachusetts 01267

(Received 26 October 1995)

We prove a new result limiting the amount of accessible information in a quantum channel. This generalizes Kholevo's theorem and implies it as a simple corollary. Our proof uses the strong subadditivity of the von Neumann entropy functional $S(\rho)$ and a specific physical analysis of the measurement process. The result presented here has application in information obtained from "weak" measurements, such as those sometimes considered in quantum cryptography. [S0031-9007(96)00060-9]

PACS numbers: 89.70.+c, 03.65.Db

When a measurement is performed on a quantum system, information is acquired about the preparation of the system. In a quantum communication channel, for example, the receiver uses a "decoding observable" to infer something about the "signal state" of the channel [1]. One of the central problems of quantum information theory is the extent to which the laws of quantum mechanics provide limitations and opportunities for information acquisition in various contexts.

A theorem stated by Levitin [2] and proved by Kholevo [3] provides an upper bound on the amount of information that may be obtained about the preparation of a quantum system by the measurement of any observable. Let Q be a quantum system which is prepared in the mixed state ρ_k with *a priori* probability p_k . The overall ensemble of states is just

$$\rho = \sum_k p_k \rho_k.$$

The information acquired in a measurement is represented by the mutual information $I(A:K)$ between the measurement outcome A and the preparation K . The mutual information is given by

$$I(A:K) = H(K) - H(K|A),$$

that is, the difference between the *a priori* uncertainty in K [measured by the Shannon entropy $H(K) = -\sum_k p_k \ln p_k$] and the average uncertainty in K after the measurement outcome A is known. Kholevo's theorem states that

$$I(A:K) \leq S(\rho) - \sum_k p_k S(\rho_k), \quad (1)$$

where $S(\rho) = -\text{Tr} \rho \ln \rho$, the von Neumann entropy of the density operator ρ . The quantity on the right-hand side of the inequality,

$$\chi = S(\rho) - \sum_k p_k S(\rho_k), \quad (2)$$

bounds the accessible information in the quantum ensemble. χ and its properties will be of central importance in this paper.

The standard proof of Kholevo's theorem is fairly difficult (see [3,4]). In this paper we provide an alternative

deviation that is better in several respects: (1) The proof is based on a well-known property of the entropy functional $S(\rho)$ known as *strong subadditivity*. [This property is not itself easy to prove, but it is a property of $S(\rho)$ itself, without reference to measurements.] (2) We employ a physical model of the measurement process instead of a mathematically defined "observable." This model is sufficient to describe any "generalized" measurement (or POVM) [5], where distinct outcomes are represented by positive operators instead of projections. (3) Because our model of measurement includes the effect of the measurement process on the system Q , we are in fact able to arrive at a stronger result than Eq. (1). Our version is much sharper than Kholevo's theorem in the case of "weak" measurements that acquire only a small fraction of the available information.

Strong subadditivity.—Suppose X and Y are a pair of quantum systems whose joint state is given by the density matrix $\rho^{(XY)}$. The individual systems are described by states $\rho^{(X)}$ and $\rho^{(Y)}$, which are given by partial traces of the joint state:

$$\begin{aligned} \rho^{(X)} &= \text{Tr}_Y \rho^{(XY)}, \\ \rho^{(Y)} &= \text{Tr}_X \rho^{(XY)}. \end{aligned}$$

Subadditivity [6] is a property of the entropy functional that states that

$$S(\rho^{(XY)}) \leq S(\rho^{(X)}) + S(\rho^{(Y)}). \quad (3)$$

A stronger property of entropy called *strong subadditivity* [7] is clearly related. Suppose X , Y , and Z are three quantum systems. Then

$$S(\rho^{(XYZ)}) + S(\rho^{(X)}) \leq S(\rho^{(XY)}) + S(\rho^{(XZ)}), \quad (4)$$

where $\rho^{(X)}$, $\rho^{(XY)}$, etc. are states of various subsystems obtained by partial traces of the global state $\rho^{(XYZ)}$. This is a nontrivial property to establish for the von Neumann entropy S ; see, for example, Wehrl's review paper on the properties of entropy [6]. The strong subadditivity of S

implies, among other things, that S is subadditive; simply consider a state of the form

$$\rho^{(XYZ)} = |\psi^{(X)}\rangle\langle\psi^{(X)}| \otimes \rho^{(YZ)}$$

for a pure state $|\psi^{(X)}\rangle$ of X and an arbitrary joint state $\rho^{(YZ)}$ of systems Y and Z .

Our interest in strong subadditivity is in proving an important property of the Kholevo bound functional χ , defined in Eq. (2) above. Suppose the joint state $\rho_k^{(XY)}$ of systems X and Y is produced with probability p_k . Then $\rho^{(XY)} = \sum_k p_k \rho_k^{(XY)}$, and

$$\chi^{(XY)} = S(\rho^{(XY)}) - \sum_k p_k S(\rho_k^{(XY)}).$$

We can ignore system Y entirely, and consider the states $\rho_k^{(X)}$ of system X alone. Then the Kholevo bound would be

$$\chi^{(X)} = S(\rho^{(X)}) - \sum_k p_k S(\rho_k^{(X)}).$$

We now wish to establish that $\chi^{(X)} \leq \chi^{(XY)}$. This is a sensible inequality; since χ is the upper bound on the accessible information, it is reasonable that this bound does not increase when part of the system is discarded. Our model of the measurement process, described in the next section, includes a unitary evolution (during which χ is constant) and a discarding of the “environment” degrees of freedom. The fact that χ cannot increase during such a process will be the key to our main result.

The proof that χ cannot increase when part of the system is discarded is based on the following fact. Let A and B be two systems in a state of the form

$$\rho^{(AB)} = \sum_k q_i |a_i^{(A)}\rangle\langle a_i^{(A)}| \otimes \rho_i^{(B)},$$

where the states $|a_i^{(A)}\rangle$ are orthogonal. Then the entropy of the joint state is

$$S(\rho^{(AB)}) = H(\vec{q}) + \sum_i q_i S(\rho_i^{(B)}).$$

The quantity $H(\vec{q})$ is the classical (Shannon) entropy of the probability distribution q_i .

Consider the state

$$\rho^{(XYZ)} = \sum_k p_k \rho_k^{(XY)} \otimes |k^{(Z)}\rangle\langle k^{(Z)}| \quad (5)$$

for an orthogonal set of states $|k^{(Z)}\rangle$. Various terms in the strong subadditivity relation [Eq. (4)] can be calculated:

$$S(\rho^{(XYZ)}) = H(\vec{p}) + \sum_k p_k S(\rho_k^{(XY)}),$$

$$S(\rho^{(XZ)}) = H(\vec{p}) + \sum_k p_k S(\rho_k^{(X)}).$$

Strong subadditivity therefore yields

$$\begin{aligned} H(\vec{p}) + \sum_k p_k S(\rho_k^{(XY)}) + S(\rho^{(X)}) \\ \leq H(\vec{p}) + \sum_k p_k S(\rho_k^{(X)}) + S(\rho^{(XY)}), \\ S(\rho^{(X)}) - \sum_k p_k S(\rho_k^{(X)}) \leq S(\rho^{(XY)}) - \sum_k p_k S(\rho_k^{(XY)}), \\ \chi^{(X)} \leq \chi^{(XY)}, \end{aligned} \quad (6)$$

which is what we wished to prove.

Measurement.—Our discussion of measurement will be based on a specific physical model of measurement, to which we now turn. Suppose we have a quantum system Q with an initial state $\rho^{(Q)}$. The measurement process will involve two additional quantum systems: an apparatus system A and an environment system E . The systems A and E are initially in a joint state $\rho_0^{(AE)}$, so that the overall initial state of everything is $\rho^{(AEQ)} = \rho_0^{(AE)} \otimes \rho^{(Q)}$. The measurement process proceeds in two successive stages. (1) A dynamical evolution including interactions among A , E , and Q , represented by a unitary operator U :

$$\rho^{(AEQ)} \rightarrow \hat{\rho}^{(AEQ)} = U \rho^{(AEQ)} U^\dagger.$$

(2) Discarding of the environment, represented by a partial trace over the system E :

$$\hat{\rho}^{(AEQ)} \rightarrow \hat{\rho}^{(AQ)} = \text{Tr}_E \hat{\rho}^{(AEQ)}.$$

For the process to constitute a measurement, we require that, after these two stages, the state $\hat{\rho}^{(AQ)}$ be of the following form:

$$\hat{\rho}^{(AQ)} = \sum_a P(a) |\phi_a^{(A)}\rangle\langle\phi_a^{(A)}| \otimes w_a^{(Q)}, \quad (7)$$

where the states $|\phi_a^{(A)}\rangle$ are a fixed orthogonal set of apparatus states, independent of the input state $\rho^{(Q)}$. [$P(a)$ and $w_a^{(Q)}$ may depend on $\rho^{(Q)}$.]

We note several things about this model. The states $|\phi_a^{(A)}\rangle$ are interpreted as the “pointer states” of the measurement apparatus, which label measurement outcomes. For ordinary measurement apparatus, these states are macroscopically distinguishable and therefore orthogonal. The choice of pointer states is not determined by the state of the system being measured, but is fixed for a given measurement device in a given environment.

It may be objected that it is unrealistic to suppose that a given measurement outcome corresponds to a *pure* state of the apparatus system A , which may after all be macroscopic. But A represents only the “pointer” degree of freedom of the apparatus; other degrees of freedom (including thermal degrees of freedom) may be considered to be part of the environment E . Furthermore, if the pointer state is not completely resolved by an outside observer, that observer will simply “bin” various values of a together, resulting in a decrease in the acquired information.

Coherences between different measurement outcomes do not remain in the joint state of systems A and Q . Any such coherences have “leaked away” into the environment during the dynamical evolution. The pointer states are in fact determined by the requirement that different measurement outcomes decohere via interaction with the environment [8].

The numbers $P(a)$ are interpreted as the probabilities of the various outcomes of the measurement, indexed by a . The states $w_a^{(Q)}$ are the possible states of Q after the measurement.

The model we have described for measurement seems very general. Any interaction of the measuring apparatus and the environment with the system Q is presumably described by unitary dynamics, and coherences between measurement outcomes are then lost. The different apparatus readings are distinguishable and therefore orthogonal. We therefore adopt this as a general picture applicable to any measurement.

Information acquisition.—As before, we suppose that the system Q is initially prepared in the state $\rho_k^{(Q)}$ with *a priori* probability p_k . The state of the apparatus A and environment E is $\rho_0^{(AE)}$, independent of the preparation of Q . (The independence assumption simply says that no additional information about the preparation of the system is available except that which resides in Q itself.) The initial state of the entire system given the k th preparation for Q is

$$\rho_k^{(AEQ)} = \rho_0^{(AE)} \otimes \rho_k^{(Q)}.$$

Averaging over the possible preparations, we obtain

$$\rho^{(AEQ)} = \sum_k p_k \rho_k^{(AEQ)}.$$

Because $\rho_0^{(AE)}$ is independent of the preparation k , $\chi^{(AEQ)} = \chi^{(Q)}$.

A measurement process as described in the previous section now takes place. In the first stage of the measurement process, the states $\rho_k^{(AEQ)}$ evolve unitarily into the states $\hat{\rho}_k^{(AEQ)}$. In the second stage, $\hat{\rho}_k^{(AEQ)}$ is reduced to $\hat{\rho}_k^{(AQ)}$ by partial trace. These states are of the form

$$\hat{\rho}_k^{(AQ)} = \sum_a P(a|k) |\phi_a^{(A)}\rangle \langle \phi_a^{(A)}| \otimes w_{ak}^{(Q)}. \quad (8)$$

$P(a|k)$ is the conditional probability for the measurement outcome a given the k th preparation for Q , and the states $w_{ak}^{(Q)}$ are the resulting states of Q when the k th preparation leads to the measurement outcome a . [The $w_{ak}^{(Q)}$ need only be defined when $P(a|k) \neq 0$.] The average state $\hat{\rho}^{(AQ)}$ of A and Q after the measurement process is

$$\hat{\rho}^{(AQ)} = \sum_{a,k} P(a,k) |\phi_a^{(A)}\rangle \langle \phi_a^{(A)}| \otimes w_{ak}^{(Q)},$$

where $P(a,k) = p_k P(a|k)$, the joint probability of measurement outcome a and preparation k . We can do the k

sum and write this as

$$\hat{\rho}^{(AQ)} = \sum_a P(a) |\phi_a^{(A)}\rangle \langle \phi_a^{(A)}| \otimes w_a^{(Q)}, \quad (9)$$

where $P(a) = \sum_k P(a,k)$ is the total probability for the measurement outcome a and

$$w_a^{(Q)} = \sum_k P(k|a) w_{ak}^{(Q)}$$

is the final state of Q , averaged over preparations, given the outcome a . [Of course, $P(k|a) = P(a,k)/P(a)$.]

$\hat{\chi}^{(AEQ)}$ must be the same as $\chi^{(AEQ)}$, since entropies are preserved under unitary evolution. But because χ is nonincreasing under the partial trace operation [Eq. (6) above], we know that $\hat{\chi}^{(AQ)} \leq \hat{\chi}^{(AEQ)}$. Thus, we conclude that

$$\hat{\chi}^{(AQ)} \leq \chi^{(Q)}. \quad (10)$$

We now evaluate $\hat{\chi}^{(AQ)} = S(\hat{\rho}^{(AQ)}) - \sum_k p_k S(\hat{\rho}_k^{(AQ)})$. The form of $\hat{\rho}^{(AQ)}$ given in Eq. (9) yields that

$$S(\hat{\rho}^{(AQ)}) = H(A) + \sum_a P(a) S(w_a^{(Q)}). \quad (11)$$

where $H(A) = -\sum_a P(a) \ln P(a)$, the Shannon entropy of the (average) probability distribution $P(a)$ over the measurement outcomes. Similarly, given Eq. (8) for $\hat{\rho}_k^{(AQ)}$,

$$S(\hat{\rho}_k^{(AQ)}) = H(A|k) + \sum_a P(a|k) S(w_{ak}^{(Q)}), \quad (12)$$

where $H(A|k)$ is the Shannon entropy of the measurement outcome conditional on a particular preparation k . The conditional information $H(A|K) = \sum_k p_k H(A|k)$.

We therefore have

$$\begin{aligned} \chi^{(Q)} &\geq S(\hat{\rho}^{(AQ)}) - \sum_k p_k S(\hat{\rho}_k^{(AQ)}) \\ &= H(A) + \sum_a P(a) S(w_a^{(Q)}) - H(A|K) \\ &\quad - \sum_{a,k} P(a,k) S(w_{ak}^{(Q)}). \end{aligned}$$

The mutual information $I(A:K) = H(A) - H(A|K)$. We can rewrite this inequality as

$$\begin{aligned} \chi^{(Q)} &\geq I(A:K) + \sum_a P(a) \\ &\quad \times \left[S(w_a^{(Q)}) - \sum_k P(k|a) S(w_{ak}^{(Q)}) \right]. \end{aligned}$$

The quantity in square brackets is $\chi_a^{(Q)}$, the value of χ for the system Q after the measurement is concluded, conditional on the outcome a for the experiment. We have therefore shown that

$$I(A:K) \leq \chi^{(Q)} - \sum_a P(a) \chi_a^{(Q)}. \quad (13)$$

Equation (13) is our central result. Since the quantities $\chi_a^{(Q)}$ are non-negative, Kholevo's theorem [Eq. (1)] is a corollary to this more general theorem.

The information obtained about the preparation of the system Q by means of a measurement procedure is thus bounded by the average amount that the quantity χ decreases in the course of the measurement. This fact expresses a new and potentially useful relationship between the *power* of the measurement to provide information about Q and the physical *effect* of the measurement process upon Q .

Suppose, for example, that our measurement is described by a set of one-dimensional projections π_a , and that the effect of the procedure is given by the "projection postulate." That is, if the outcome a is obtained for an input state $\rho^{(Q)}$, then the final state is simply π_a . Then $\chi_a^{(Q)} = 0$ for all a . In other words, the final state of the system Q depends only upon the measurement outcome and not upon the preparation (except inasmuch as the preparation determines the probabilities of the various outcomes). In this case, we simply have $I(A:K) \leq \chi^{(Q)}$.

However, if the measurement is not complete, and the "outcome operators" π_a are projections onto subspaces of many dimensions, then the natural generalization of the projection postulate yields final states (given input $\rho^{(Q)}$)

$$w_a^{(Q)} = \frac{\pi_a \rho^{(Q)} \pi_a}{\text{Tr} \pi_a \rho^{(Q)} \pi_a}.$$

These may depend upon the preparation directly, and thus Eq. (13) yields a stronger bound for $I(A:K)$ than Kholevo's theorem. Another example where Eq. (13) may improve upon the Kholevo bound is the "translucent" measurements [9] that have been studied in connection with eavesdropping in quantum cryptography

[10]. Whenever a measurement "leaves some information behind" in the system Q , Eq. (13) will in general be a stronger statement about the information $I(A:K)$ that is obtained.

The authors wish to acknowledge helpful conversations with Richard Jozsa, Carlton Caves, Chris Fuchs, and Howard Barnum.

-
- [1] B. W. Schumacher, in *Complexity, Entropy, and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Redwood City, CA, 1990), pp. 29–37.
 - [2] L. B. Levitin, in *Proceedings of the Fourth All-Union Conference on Information and Coding Theory* (Mockva-Tashkent, Tashkent, 1969), Sec. II; also L. B. Levitin, in *Information Complexity and Control in Quantum Physics*, edited by A. Blaquiéve, S. Diner, and G. Lochak (Springer, New York, 1987), pp. 15–47.
 - [3] A. S. Kholevo, *Probl. Peredachi Inf.* **9**, 3 (1973) [*Probl. Inf. Transm. (USSR)* **9**, 177 (1973)].
 - [4] C. Caves and C. Fuchs, *Phys. Rev. Lett.* **73**, 3047 (1994).
 - [5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976), pp. 74–83.
 - [6] A. Wehrl, *Rev. Mod. Phys.* **50**, 221 (1978).
 - [7] E. H. Lieb and M. B. Ruskai, *J. Math. Phys. (N.Y.)* **14**, 267 (1973).
 - [8] W. H. Zurek, *Phys. Rev. D* **26**, 1862 (1982).
 - [9] A. Ekert, B. Huttner, M. Palma, and A. Peres, *Phys. Rev. A* **50**, 1047 (1994).
 - [10] C. H. Bennett and G. Brassard, *Proceedings of the IEEE Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 175; C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, *J. Cryptology* **5**, 3 (1992).