

# The Entropy Photon-Number Inequality and its Consequences\*

Saikat Guha, Baris I. Erkmen, and Jeffrey H. Shapiro  
*Massachusetts Institute of Technology, Research Laboratory of Electronics*

Determining the ultimate classical information carrying capacity of electromagnetic waves requires quantum-mechanical analysis to properly account for the bosonic nature of these waves. Recent work has established capacity theorems for bosonic single-user, broadcast, and wiretap channels, under the presumption of two minimum output entropy conjectures. Despite considerable accumulated evidence that supports the validity of these conjectures, they have yet to be proven. Here we show that the preceding minimum output entropy conjectures are simple consequences of an Entropy Photon-Number Inequality, which is a conjectured quantum-mechanical analog of the Entropy Power Inequality from classical information theory.

## I. MOTIVATION AND HISTORY

The performance of communication systems that rely on electromagnetic wave propagation are ultimately limited by noise of quantum-mechanical origin. Moreover, high-sensitivity photodetection systems have long been close to or at this noise limit, hence determining the ultimate capacities of lasercom channels, which requires quantum-mechanical analysis, is of immediate relevance. The most famous channel capacity formula is Shannon's result for the classical additive white Gaussian noise channel. For a complex-valued channel model in which we transmit  $a$  and receive  $c = \sqrt{\eta}a + \sqrt{1-\eta}b$ , where  $0 < \eta < 1$  is the channel's transmissivity and  $b$  is a zero-mean, isotropic, complex-valued Gaussian random variable that is independent of  $a$ , Shannon's capacity is

$$C_{\text{classical}} = \ln[1 + \eta\bar{N}/(1-\eta)N] \text{ nats/use}, \quad (1)$$

with  $E(|a|^2) \leq \bar{N}$  and  $E(|b|^2) = N$ . In the quantum version of this channel model, we control the state of an electromagnetic mode with photon annihilation operator  $\hat{a}$  at the transmitter, and receive another mode with photon annihilation operator  $\hat{c} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{b}$ , where  $\hat{b}$  is the annihilation operator of a noise mode that is in a zero-mean, isotropic, complex-valued Gaussian state. For lasercom, if quantum measurements corresponding to ideal optical homodyne or heterodyne detection are employed at the receiver, this quantum channel reduces to a real-valued (homodyne) or complex-valued (heterodyne) additive Gaussian noise channel, from which the following capacity formulas (in nats/use) follow:

$$C_{\text{homodyne}} = \frac{1}{2} \ln[1 + 4\eta\bar{N}/(2(1-\eta)N + 1)] \quad (2)$$

$$C_{\text{heterodyne}} = \ln[1 + \eta\bar{N}/((1-\eta)N + 1)]. \quad (3)$$

The + 1 terms in the capacity formulas' noise denominators are quantum noise contributions, so that even when

the noise mode  $\hat{b}$  is unexcited—in its vacuum state—these capacities remain finite, unlike the situation in Eq. (1).

The classical capacity of the pure-loss bosonic channel—in which the  $\hat{b}$  mode is unexcited ( $N = 0$ )—was shown in [1] to be  $C_{\text{pure-loss}} = g(\eta\bar{N})$  nats/use, where  $g(x) \equiv (x+1)\ln(x+1) - x\ln(x)$  is the Shannon entropy of the Bose-Einstein probability distribution with mean  $x$ . This capacity exceeds the  $N = 0$  versions of Eqs. (2) and (3), as well as the best known bound on the capacity of photon-number measurement (ideal optical direct detection). The ultimate capacity of the thermal-noise ( $N > 0$ ) version of this channel was considered in [2], where the lower bound  $C_{\text{thermal}} \geq g(\eta\bar{N} + (1-\eta)N) - g((1-\eta)N)$  was shown to be the capacity if the thermal channel obeyed a certain minimum output entropy conjecture. This conjecture states that the von Neumann entropy at the output of the thermal channel is minimized when the  $\hat{a}$  mode is in its vacuum state. Several partial results, numerical evidence, and a suite of upper and lower bounds were obtained to support the conjecture [3], but it has yet to be proven. Nevertheless, the preceding lower bound already exceeds Eqs. (2) and (3) as well as the best known bounds on the capacity of direct detection.

The thermal channel work was followed by a capacity analysis of the bosonic broadcast channel, for which we obtained an inner bound on the capacity region [4], which we showed to be the capacity region under the presumption of a second minimum output entropy conjecture that was a dual to the first. Both conjectures have been proven if the input states are restricted to be Gaussian, and we have shown that they are equivalent under this input-state restriction. In unpublished work [5], we have shown that the second conjecture will establish the privacy capacity of the lossy bosonic channel, as well as its ultimate quantum information carrying capacity.

The Entropy Power Inequality (EPI) from classical information theory is widely used in coding theorem converse proofs for Gaussian channels. By analogy with the EPI, we conjecture its quantum version, viz., the Entropy Photon-number Inequality (EPnI). In this paper we show that the two minimum output entropy conjectures cited above are simple corollaries of the EPnI. Hence, proving the EPnI would immediately establish key results for the capacities of bosonic communication channels.

---

\*Research supported by the Defense Advanced Research Projects Agency and by the W. M. Keck Foundation Center for Extreme Quantum Information Theory.

## II. DESCRIPTION OF THE PROBLEM

### A. The Entropy Power Inequality

Let  $\mathbf{X}$  and  $\mathbf{Y}$  be statistically independent,  $n$ -dimensional, real-valued random vectors that possess differential (Shannon) entropies  $h(\mathbf{X})$  and  $h(\mathbf{Y})$  respectively. Because a real-valued, zero-mean Gaussian random variable  $U$  has differential entropy given by  $h(U) = \ln(2\pi e \langle U^2 \rangle)$ , where the mean-squared value,  $\langle U^2 \rangle$ , is considered to be the *power* of  $U$ , the entropy powers of  $\mathbf{X}$  and  $\mathbf{Y}$  are taken to be

$$P(\mathbf{X}) \equiv \frac{e^{h(\mathbf{X})/n}}{2\pi e} \quad \text{and} \quad P(\mathbf{Y}) \equiv \frac{e^{h(\mathbf{Y})/n}}{2\pi e}. \quad (4)$$

In this way, an  $n$ -dimensional, real-valued, random vector  $\tilde{\mathbf{X}}$  comprised of independent, identically distributed (i.i.d.), real-valued, zero-mean, variance- $P(\mathbf{X})$ , Gaussian random variables has differential entropy  $h(\tilde{\mathbf{X}}) = h(\mathbf{X})$ . We can similarly define an i.i.d. Gaussian random vector  $\tilde{\mathbf{Y}}$  with differential entropy  $h(\tilde{\mathbf{Y}}) = h(\mathbf{Y})$ . We define a new random vector by

$$\mathbf{Z} \equiv \sqrt{\eta} \mathbf{X} + \sqrt{1-\eta} \mathbf{Y}, \quad (5)$$

where  $0 \leq \eta \leq 1$ . This random vector has differential entropy  $h(\mathbf{Z})$  and entropy power  $P(\mathbf{Z})$ . Furthermore, let  $\tilde{\mathbf{Z}} \equiv \sqrt{\eta} \tilde{\mathbf{X}} + \sqrt{1-\eta} \tilde{\mathbf{Y}}$ . Three equivalent forms of the Entropy Power Inequality (EPI), see, e.g., [6], are then:

$$P(\mathbf{Z}) \geq \eta P(\mathbf{X}) + (1-\eta)P(\mathbf{Y}) \quad (6)$$

$$h(\mathbf{Z}) \geq h(\tilde{\mathbf{Z}}) \quad (7)$$

$$h(\mathbf{Z}) \geq \eta h(\mathbf{X}) + (1-\eta)h(\mathbf{Y}). \quad (8)$$

### B. The Entropy Photon-Number Inequality

Let  $\hat{\mathbf{a}} = [\hat{a}_1 \hat{a}_2 \cdots \hat{a}_n]$  and  $\hat{\mathbf{b}} = [\hat{b}_1 \hat{b}_2 \cdots \hat{b}_n]$  be vectors of photon annihilation operators for a collection of  $2n$  different electromagnetic field modes of frequency  $\omega$  [7]. The joint state of the modes associated with  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$  is assumed to be the product-state density operator  $\hat{\rho}_{\mathbf{ab}} = \hat{\rho}_{\mathbf{a}} \otimes \hat{\rho}_{\mathbf{b}}$ , where  $\hat{\rho}_{\mathbf{a}}$  and  $\hat{\rho}_{\mathbf{b}}$  are the density operators associated with the  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$  modes. The von Neumann entropies of the  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$  modes are  $S(\hat{\rho}_{\mathbf{a}}) = -\text{tr}[\hat{\rho}_{\mathbf{a}} \ln(\hat{\rho}_{\mathbf{a}})]$  and  $S(\hat{\rho}_{\mathbf{b}}) = -\text{tr}[\hat{\rho}_{\mathbf{b}} \ln(\hat{\rho}_{\mathbf{b}})]$  [8].

The thermal state of a mode with annihilation operator  $\hat{a}$  has two equivalent forms

$$\hat{\rho}_T = \int d^2\alpha \frac{e^{-|\alpha|^2/N}}{\pi N} |\alpha\rangle\langle\alpha|, \quad (9)$$

and

$$\hat{\rho}_T = \sum_{i=0}^{\infty} \frac{N^i}{(N+1)^{i+1}} |i\rangle\langle i|, \quad (10)$$

where  $N = \langle \hat{a}^\dagger \hat{a} \rangle$  is the average photon number. In Eq. (9),  $|\alpha\rangle$  is the coherent state of amplitude  $\alpha$ , i.e., it satisfies  $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$ , for  $\alpha$  a complex number. In Eq. (10),  $|i\rangle$  is the  $i$ -photon state, i.e., it satisfies  $\hat{N}|i\rangle = i|i\rangle$ , for  $i = 0, 1, 2, \dots$ , with  $\hat{N} \equiv \hat{a}^\dagger \hat{a}$  being the photon number operator. Physically, Eq. (9) says that the thermal state is an isotropic Gaussian mixture of coherent states. Equation (10), on the other hand, says that the thermal state is a Bose-Einstein mixture of number states. From Eq. (10) we immediately have that  $S(\hat{\rho}_T) = g(N)$ , because the photon-number states are orthonormal [9]. Note that  $g(N)$ , for  $N \geq 0$ , is a non-negative, monotonically increasing, concave function of  $N$ . Thus it has an inverse,  $g^{-1}(S)$ , for  $S \geq 0$ , that is non-negative, monotonically increasing, and convex.

The entropy photon-numbers of the density operators  $\hat{\rho}_{\mathbf{a}}$  and  $\hat{\rho}_{\mathbf{b}}$  are defined as follows:

$$N(\hat{\rho}_{\mathbf{a}}) \equiv g^{-1}(S(\hat{\rho}_{\mathbf{a}})/n) \quad \text{and} \quad N(\hat{\rho}_{\mathbf{b}}) \equiv g^{-1}(S(\hat{\rho}_{\mathbf{b}})/n). \quad (11)$$

Thus, if  $\hat{\rho}_{\hat{\mathbf{a}}} \equiv \bigotimes_{i=1}^n \hat{\rho}_{T_{a_i}}$  and  $\hat{\rho}_{\hat{\mathbf{b}}} \equiv \bigotimes_{i=1}^n \hat{\rho}_{T_{b_i}}$ , where  $\hat{\rho}_{T_{a_i}}$  is the thermal state of average photon number  $N(\hat{\rho}_{\mathbf{a}})$  for the  $\hat{a}_i$  mode and  $\hat{\rho}_{T_{b_i}}$  is the thermal state of average photon number  $N(\hat{\rho}_{\mathbf{b}})$  for the  $\hat{b}_i$  mode, then we have  $S(\hat{\rho}_{\hat{\mathbf{a}}}) = S(\hat{\rho}_{\mathbf{a}})$  and  $S(\hat{\rho}_{\hat{\mathbf{b}}}) = S(\hat{\rho}_{\mathbf{b}})$ . We define a new vector of photon annihilation operators,  $\hat{\mathbf{c}} = [\hat{c}_1 \hat{c}_2 \cdots \hat{c}_n]$ , by

$$\hat{\mathbf{c}} \equiv \sqrt{\eta} \hat{\mathbf{a}} + \sqrt{1-\eta} \hat{\mathbf{b}}, \quad \text{for } 0 \leq \eta \leq 1, \quad (12)$$

and use  $\hat{\rho}_{\mathbf{c}}$  to denote its density operator. This is equivalent to saying that  $\hat{c}_i$  is the output of a lossless beam splitter whose inputs,  $\hat{a}_i$  and  $\hat{b}_i$ , couple to that output with transmissivity  $\eta$  and reflectivity  $1-\eta$ , respectively.

We can now state two equivalent forms of our conjectured Entropy Photon-Number Inequality (EPnI) [10]:

$$N(\hat{\rho}_{\mathbf{c}}) \geq \eta N(\hat{\rho}_{\mathbf{a}}) + (1-\eta)N(\hat{\rho}_{\mathbf{b}}) \quad (13)$$

$$S(\hat{\rho}_{\mathbf{c}}) \geq S(\hat{\rho}_{\hat{\mathbf{c}}}), \quad (14)$$

where  $\hat{\rho}_{\hat{\mathbf{c}}} \equiv \bigotimes_{i=1}^n \hat{\rho}_{T_{c_i}}$  with  $\hat{\rho}_{T_{c_i}}$  being the thermal state of mean photon number  $\eta N(\hat{\rho}_{\mathbf{a}}) + (1-\eta)N(\hat{\rho}_{\mathbf{b}})$  for  $\hat{c}_i$ .

## III. PRIOR WORK AND DISCUSSION

By analogy with the classical EPI, we might expect there to be a third equivalent form of the quantum EPnI, viz.,

$$S(\hat{\rho}_{\mathbf{c}}) \geq \eta S(\hat{\rho}_{\mathbf{a}}) + (1-\eta)S(\hat{\rho}_{\mathbf{b}}). \quad (15)$$

It is easily shown that Eq. (13) implies Eq. (15) [11], but we have not been able to prove the converse. Indeed, we suspect that the converse might be false. More important than whether or not (15) is equivalent to Eq. (13) and Eq. (14), is the role of the EPnI in proving classical information capacity results for bosonic channels. In particular, the EPnI provides simple proofs of the following

two minimum output entropy conjectures. These conjectures are important because proving minimum output entropy conjecture 1 also proves the conjectured capacity of the thermal-noise channel [2], and proving minimum output entropy conjecture 2 also proves the conjectured capacity region of the bosonic broadcast channel [4].

**Minimum Output Entropy Conjecture 1** — Let  $\mathbf{a}$  and  $\mathbf{b}$  be  $n$ -dimensional vectors of annihilation operators, with joint density operator  $\hat{\rho}_{\mathbf{ab}} = (|\psi\rangle_{\mathbf{aa}}\langle\psi|) \otimes \hat{\rho}_{\mathbf{b}}$ , where  $|\psi\rangle_{\mathbf{a}}$  is an arbitrary zero-mean-field pure state of the  $\mathbf{a}$  modes and  $\hat{\rho}_{\mathbf{b}} = \bigotimes_{i=1}^n \hat{\rho}_{T_{b_i}}$  with  $\hat{\rho}_{T_{b_i}}$  being the  $\hat{b}_i$  mode's thermal state of average photon number  $K$ . Define a new vector of photon annihilation operators,  $\hat{\mathbf{c}} = [\hat{c}_1 \hat{c}_2 \cdots \hat{c}_n]$ , by (12) and use  $\hat{\rho}_{\mathbf{c}}$  to denote its density operator and  $S(\hat{\rho}_{\mathbf{c}})$  to denote its von Neumann entropy. Then choosing  $|\psi\rangle_{\mathbf{a}}$  to be the  $n$ -mode vacuum state minimizes  $S(\hat{\rho}_{\mathbf{c}})$ .

**Minimum Output Entropy Conjecture 2** — Let  $\mathbf{a}$  and  $\mathbf{b}$  be  $n$ -dimensional vectors of annihilation operators, as in the previous section, with joint density operator  $\hat{\rho}_{\mathbf{ab}} = (|\psi\rangle_{\mathbf{aa}}\langle\psi|) \otimes \hat{\rho}_{\mathbf{b}}$ , where  $|\psi\rangle_{\mathbf{a}} = \bigotimes_{i=1}^n |0\rangle_{a_i}$  is the  $n$ -mode vacuum state and  $\hat{\rho}_{\mathbf{b}}$  has von Neumann entropy  $S(\hat{\rho}_{\mathbf{b}}) = ng(K)$  for some  $K \geq 0$ . Define a new vector of photon annihilation operators,  $\hat{\mathbf{c}} = [\hat{c}_1 \hat{c}_2 \cdots \hat{c}_n]$ , by (12) and use  $\hat{\rho}_{\mathbf{c}}$  to denote its density operator and  $S(\hat{\rho}_{\mathbf{c}})$  to denote its von Neumann entropy. Then choosing  $\hat{\rho}_{\mathbf{b}} = \bigotimes_{i=1}^n \hat{\rho}_{T_{b_i}}$  with  $\hat{\rho}_{T_{b_i}}$  being the  $\hat{b}_i$  mode's thermal state of average photon number  $K$  minimizes  $S(\hat{\rho}_{\mathbf{c}})$ .

To see that the EPnI encompasses both of the preceding minimum output entropy conjectures is our final task in this paper. We begin by using the premise of conjecture 1 in Eq. (13). Because the  $\hat{\mathbf{a}}$  modes are in a pure state, we get  $S(\hat{\rho}_{\mathbf{a}}) = 0$  and hence the EPnI tells us that

$$N(\hat{\rho}_{\mathbf{c}}) \geq (1 - \eta)N(\hat{\rho}_{\mathbf{b}}) = (1 - \eta)K. \quad (16)$$

Taking  $g(\cdot)$  on both sides of this inequality, we get  $S(\hat{\rho}_{\mathbf{c}})/n \geq g[(1 - \eta)K]$ . But, if  $|\psi\rangle_{\mathbf{a}}$  is the  $n$ -mode vacuum state, we can easily show that  $\hat{\rho}_{\mathbf{c}} = \bigotimes_{i=1}^n \hat{\rho}_{T_{c_i}}$ , with  $\hat{\rho}_{T_{c_i}}$  being the  $\hat{c}_i$  mode's thermal state of average photon number  $(1 - \eta)K$ . Thus, when  $|\psi\rangle_{\mathbf{a}}$  is the  $n$ -mode vacuum state we get  $S(\hat{\rho}_{\mathbf{c}}) = ng[(1 - \eta)K]$ , which completes the proof.

Next, we apply the premise of conjecture 2 in Eq. (13). Once again, the  $\hat{\mathbf{a}}$  modes are in a pure state, so we get

$$N(\hat{\rho}_{\mathbf{c}}) \geq (1 - \eta)N(\hat{\rho}_{\mathbf{b}}) = (1 - \eta)K, \quad (17)$$

and hence  $S(\hat{\rho}_{\mathbf{c}})/n \geq g[(1 - \eta)K]$ . But, taking  $\hat{\rho}_{\mathbf{b}} = \bigotimes_{i=1}^n \hat{\rho}_{T_{b_i}}$ , with  $\hat{\rho}_{T_{b_i}}$  being the  $\hat{b}_i$  mode's thermal state of average photon number  $K$ , satisfies the premise of minimum output entropy conjecture 2 and implies that  $\hat{\rho}_{\mathbf{c}} = \bigotimes_{i=1}^n \hat{\rho}_{T_{c_i}}$ , with  $\hat{\rho}_{T_{c_i}}$  being the  $\hat{c}_i$  mode's thermal state of average photon number  $(1 - \eta)K$ . In this case we have  $S(\hat{\rho}_{\mathbf{c}}) = ng[(1 - \eta)K]$ , which completes the proof.

- 
- [1] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, Phys. Rev. Lett. **92**, 027902 (2004).  
 [2] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, B. J. Yen, and H. P. Yuen, in *Quantum Information, Statistics, Probability*, edited by O. Hirota (Rinton Press, Princeton, NJ, 2004), pp. 90–101.  
 [3] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, and J. H. Shapiro, Phys. Rev. A **70**, 032315 (2004).  
 [4] S. Guha, J. H. Shapiro, and B. I. Erkmen, Phys. Rev. A **76**, 032303 (2007).  
 [5] S. Guha, J. H. Shapiro, and B. I. Erkmen, arXiv quant-ph/0801.0841  
 [6] O. Rioul, “Information theoretic proofs of entropy power inequalities,” arXiv cs.IT/0704.175  
 [7] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, (Cambridge University Press, Cambridge, 1995).  
 [8] A density operator is Hermitian, with eigenvalues that form a probability distribution. Thus, the von Neumann entropy of a density operator  $\hat{\rho}$  is the Shannon entropy of its eigenvalues.  
 [9] The coherent states,  $\{|\alpha\rangle\}$ , are *not* orthonormal, but rather overcomplete.  
 [10] To show that (13) implies (14), assume (13) is true:

$$N(\hat{\rho}_{\mathbf{c}}) \geq \eta N(\hat{\rho}_{\mathbf{a}}) + (1 - \eta)N(\hat{\rho}_{\mathbf{b}}) \quad (18)$$

$$= \eta N(\hat{\rho}_{\hat{\mathbf{a}}}) + (1 - \eta)N(\hat{\rho}_{\hat{\mathbf{b}}}) \quad (19)$$

Now, if  $\hat{\rho}_{\hat{\mathbf{a}}\hat{\mathbf{b}}} = \hat{\rho}_{\hat{\mathbf{a}}} \otimes \hat{\rho}_{\hat{\mathbf{b}}}$  is the joint density operator of the  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$  modes, we find that the state of the  $\hat{\mathbf{c}}$  modes

is  $\hat{\rho}_{\hat{\mathbf{c}}} \equiv \bigotimes_{i=1}^n \hat{\rho}_{T_{c_i}}$ , where  $\hat{\rho}_{T_{c_i}}$  is a thermal state with average photon number given by  $N(\hat{\rho}_{\hat{\mathbf{c}}}) = \eta N(\hat{\rho}_{\hat{\mathbf{a}}}) + (1 - \eta)N(\hat{\rho}_{\hat{\mathbf{b}}})$ , so that  $S(\hat{\rho}_{\hat{\mathbf{c}}}) = ng[N(\hat{\rho}_{\hat{\mathbf{c}}})]$ . Thus, from (19) we get  $N(\hat{\rho}_{\mathbf{c}}) \geq N(\hat{\rho}_{\hat{\mathbf{c}}}) = g^{-1}(S(\hat{\rho}_{\hat{\mathbf{c}}})/n)$ . Taking  $g(\cdot)$  of both sides of this inequality completes the proof.

To show that (14) implies (13), assume (14) is true:

$$\begin{aligned} N(\hat{\rho}_{\mathbf{c}}) &= g^{-1}(S(\hat{\rho}_{\mathbf{c}})/n) \\ &\geq g^{-1}(S(\hat{\rho}_{\hat{\mathbf{c}}})/n) = g^{-1}[g(\eta N(\hat{\rho}_{\hat{\mathbf{a}}}) + (1 - \eta)N(\hat{\rho}_{\hat{\mathbf{b}}}))] \\ &= \eta N(\hat{\rho}_{\hat{\mathbf{a}}}) + (1 - \eta)N(\hat{\rho}_{\hat{\mathbf{b}}}) \\ &= \eta N(\hat{\rho}_{\mathbf{a}}) + (1 - \eta)N(\hat{\rho}_{\mathbf{b}}), \end{aligned} \quad (20)$$

where the inequality is due to  $g^{-1}(S)$  being a monotonically increasing function of  $S$ , and the proof is complete.

- [11] Assume that (13) is true. We then have that  $N(\hat{\rho}_{\mathbf{c}}) \geq \eta N(\hat{\rho}_{\mathbf{a}}) + (1 - \eta)N(\hat{\rho}_{\mathbf{b}})$ , so that

$$S(\hat{\rho}_{\mathbf{c}}) = ng[N(\hat{\rho}_{\mathbf{c}})] \geq ng[\eta N(\hat{\rho}_{\mathbf{a}}) + (1 - \eta)N(\hat{\rho}_{\mathbf{b}})] \quad (21)$$

$$\geq \eta ng[N(\hat{\rho}_{\mathbf{a}})] + (1 - \eta)ng[N(\hat{\rho}_{\mathbf{b}})] \quad (22)$$

$$= \eta S(\hat{\rho}_{\mathbf{a}}) + (1 - \eta)S(\hat{\rho}_{\mathbf{b}}), \quad (23)$$

where the second inequality follows from  $g(N)$  being concave, and the proof is complete.