On some special cases of the Entropy Photon-Number Inequality

Smarajit Das, Naresh Sharma, and Siddharth Muthukrishnan*

July 13, 2011

Abstract

We show that the Entropy Photon-Number Inequality (EPnI) holds where one of the input states is the vacuum state and for several candidates of the other input state that includes the cases when the state has the eigenvectors as the number states and either has only two non-zero eigenvalues or has arbitrary number of non-zero eigenvalues but is a high entropy state. We also discuss the conditions, which if satisfied, would lead to an extension of these results.

1 Introduction

The Entropy Photon Number Inequality (EPnI) was conjectured by Guha et. al. [1]. EPnI has a classical analogue called Entropy power inequality which is stated as follows. Let X and Y be independent random variables with densities and $h(X)$ be the differential entropy of X, then

$$\mathrm{e}^{2h(X)} + \mathrm{e}^{2h(Y)} \leq \mathrm{e}^{2h(X+Y)} \tag{1}$$

holds. It was first stated by Shannon in Ref. [2] and the proof was given by Stam and Blachman [3, 4].

The EPnI has some important consequences in quantum information theory. In particular, if this conjecture is true, then one would be able to establish the classical capacity of certain bosonic channels [1, 5]. EPnI is shown to imply two minimum output entropy conjectures, which would suffice to prove the capacity of several other channels such as the thermal noise channel [5] and the bosonic broadcast channel [6, 7].

The statement of the inequality is as follows. Let a and b be the photon annihilation operators and let the joint state of the modes associated with a and b be the product state, i.e., $\rho_{AB} = \rho_A \otimes \rho_B$, where ρ_A and ρ_B are the density operators associated with the a and b modes respectively. For the beam-splitter with inputs a and b and output c with transmissivity η and reflectivity $1-\eta$ respectively, the annihilation operator evolution is given by

$$c = \sqrt{\eta}a + \sqrt{1-\eta}b, \tag{2}$$

*S. Das, N. Sharma and S. Muthukrishnan are with the School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai 400 005, India. email:{smarajit, nsharma}@tifr.res.in, smkrish@tcs.tifr.res.in.

The EPnI is now stated as

$$g^{-1}\left[S(\rho_C)\right] \geq \eta g^{-1}\left[S(\rho_A)\right] + (1-\eta) g^{-1}\left[S(\rho_B)\right], \tag{3}$$

where

$$g(x) = (x+1)\log(x+1) - x\log(x) \tag{4}$$

is the von Neumann entropy of the thermal state with mean photon-number x, and $S(\rho) = -\mathrm{Tr}(\rho\log\rho)$ is the von Neumann entropy.

In this paper, we prove the EPnI for the case of ρ_B to be the vacuum state, ρ_A having its eigenvectors as the number states and either having two nonzero eigenvalues or high von Neumann entropy with arbitrary number of eigenvalues. There are other candidates as well for which some special cases EPnI hold and these are mentioned later.

2 The beam-splitter transformation

We obtain the output density matrix ρ_C from the beam-splitter transformations. The annihilation operators for the two outputs are

$$c = \sqrt{\eta}a + \sqrt{1-\eta}b, \tag{5}$$

$$d = \mathrm{e}^{\iota\phi}(\sqrt{1-\eta}a - \sqrt{\eta}b), \tag{6}$$

where $[a, a^\dagger] = [b, b^\dagger] = [c, c^\dagger] = [d, d^\dagger] = \mathbb{I}$ and $[a, b] = [a, c] = [a, d] = 0$ and so on. We assume that the inputs density operators are diagonal in the number state basis and hence,

$$\rho_{AB} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} x_i y_j \left|i\right\rangle_A \left|j\right\rangle_B \left\langle i\right|_A \left\langle j\right|_B, \tag{7}$$

where x_i and y_j are the ith and jth eigenvalues of A and B respectively, $\left|i\right\rangle_A$ and $\left|j\right\rangle_B$ are the Fock number states for the systems A and B respectively. Any state $\left|i\right\rangle_A \left|j\right\rangle_B$ can be written as (see Ref. [8] for example)

$$\left|i\right\rangle_A \left|j\right\rangle_B = \frac{(a^\dagger)^i}{\sqrt{i!}}\frac{(b^\dagger)^j}{\sqrt{j!}} \left|0\right\rangle_A \left|0\right\rangle_B. \tag{8}$$

From (5) and (6), we get $a^\dagger = \sqrt{\eta}c^\dagger + \sqrt{1-\eta}\mathrm{e}^{\iota\phi}d^\dagger$ and $b^\dagger = \sqrt{1-\eta}c^\dagger - \sqrt{\eta}\mathrm{e}^{\iota\phi}d^\dagger$. Using these with (8), we get the transformation

$$\left|i\right\rangle_A \left|j\right\rangle_B \xrightarrow{\text{B.S.}} \frac{(\sqrt{\eta}c^\dagger + \sqrt{1-\eta}\mathrm{e}^{\iota\phi}d^\dagger)^i}{\sqrt{i!}}\frac{(\sqrt{1-\eta}c^\dagger - \sqrt{\eta}\mathrm{e}^{\iota\phi}d^\dagger)^j}{\sqrt{j!}} \left|0\right\rangle_C \left|0\right\rangle_D, \tag{9}$$

where B.S. indicates the action of the beam splitter. Using the fact that the operators $c^\dagger$ and $d^\dagger$ commute and the binomial expansion, we get

$$\begin{aligned}\left|i\right\rangle_A \left|j\right\rangle_B \xrightarrow{\text{B.S.}} & \frac{1}{\sqrt{i!}\sqrt{j!}}\sum_{k=0}^{i}\sum_{l=0}^{j} \mathrm{e}^{\iota(k+l)\phi}(-1)^l \binom{i}{k}\binom{j}{l}\eta^{\frac{i-k+l}{2}}(1-\eta)^{\frac{j-l+k}{2}} \\ & (c^\dagger)^{(i+j)-(k+l)}(d^\dagger)^{k+l}\left|0\right\rangle_C \left|0\right\rangle_D. \end{aligned} \tag{10}$$

Incorporating the action of $c^\dagger$ and $d^\dagger$ on the vacuum states of C and D, we get

$$|i\rangle_A |j\rangle_B \xrightarrow{\text{B.S.}} \frac{1}{\sqrt{i!}\sqrt{j!}} \sum_{k=0}^{i} \sum_{l=0}^{j} \mathrm{e}^{\iota(k+l)\phi}(-1)^l \binom{i}{k}\binom{j}{l} \eta^{\frac{i-k+l}{2}} (1-\eta)^{\frac{j-l+k}{2}} \sqrt{[(i+j)-(k+l)]!(k+l)!}\ |(i+j)-(k+l)\rangle_C\ |k+l\rangle_D \,. \tag{11}$$

Hence, we arrive at the expression for ρ_{CD} as

$$\rho_{CD} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} x_i y_j \frac{1}{i!j!} \sum_{k=0}^{i}\sum_{l=0}^{j}\sum_{k'=0}^{i}\sum_{l'=0}^{j} \mathrm{e}^{\iota[(k+l)-(k'+l')]\phi}(-1)^{l+l'}\binom{i}{k}\binom{j}{l}\binom{i}{k'}\binom{j}{l'} \eta^{i-\frac{k+k'}{2}+\frac{l+l'}{2}}(1-\eta)^{j-\frac{l+l'}{2}+\frac{k+k'}{2}} \sqrt{[(i+j)-(k+l)]!(k+l)!}\sqrt{[(i+j)-(k'+l')]!(k'+l')!} |(i+j)-(k+l)\rangle_C |k+l\rangle_D \langle (i+j)-(k'+l')|_C \langle k'+l'|_D \,. \tag{12}$$

Now, tracing out system D, we get

$$\rho_C = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} x_i y_j \frac{1}{i!j!} \sum_{k=0}^{i}\sum_{l=0}^{j}\sum_{k'=0}^{i}\sum_{l'=0}^{j} (-1)^{l+l'}\binom{i}{k}\binom{j}{l}\binom{i}{k'}\binom{j}{l'} \eta^{i-\frac{k+k'}{2}+\frac{l+l'}{2}}(1-\eta)^{j-\frac{l+l'}{2}+\frac{k+k'}{2}} [(i+j)-(k+l)]!(k+l)!\ |(i+j)-(k+l)\rangle \langle (i+j)-(k+l)|\ \delta_{k+l,k'+l'} . \tag{13}$$

We now consider the special case when ρ_B is a vacuum state. Let the set of all probability vectors (with infinite length) be denoted by $\mathbb{P}$ and if $\boldsymbol{x} \in \mathbb{P}$, then $\sum_{i=0}^{\infty} x_i = 1$ and $x_i \geq 0\ \forall\ i \geq 0$. Then (13) reduces to

$$\rho_C = \sum_{i=0}^{\infty} z_i |i\rangle_C \langle i|_C \,, \tag{14}$$

where $\boldsymbol{z} = M_\eta(\boldsymbol{x}) \triangleq M(\eta, \boldsymbol{x})$, $M : [0,1] \times \mathbb{P} \to \mathbb{P}$ is a transformation given by

$$z_i = \sum_{k=i}^{\infty} \binom{k}{i} \eta^i (1-\eta)^{k-i} x_k. \tag{15}$$

Hence, (3) reduces to

$$g^{-1}\left\{H[M_\eta(\boldsymbol{x})]\right\} \geq \eta g^{-1}\left[H(\boldsymbol{x})\right]. \tag{16}$$

Note that this equation is expected to hold for all $\boldsymbol{x} \in \mathbb{P}$ and $\eta \in [0,1]$. The inequality is trivially true for $\eta = 0$ since $M_0(\boldsymbol{x}) = [1, 0, ...]$ implying $H[M_0(\boldsymbol{x})] = 0$, and for $\eta = 1$ since $M_1(\boldsymbol{x}) = \boldsymbol{x}$.

3 ρ_A is two-dimensional in the number state basis and ρ_B is the vacuum state

Let

$$H_b(p) \triangleq -p\log(p) - (1-p)\log(1-p) \tag{17}$$

to be the binary entropy of a two-point probability distribution $[p, 1-p]$ with $0 \leq p \leq 1$. Let the eigenvalues of ρ_A given by the probability vector $\boldsymbol{x} = [1-\alpha, \alpha, 0, ...]$. Therefore, $H(\boldsymbol{x}) = H_b(\alpha)$ and $H[M_\eta(\boldsymbol{x})] = H_b(\eta\alpha)$. We now prove (16) for the above case.

Lemma 1. *For all $\eta \in [0,1]$ and $\alpha \in [0,1]$, we have*

$$g^{-1}\left[H_b(\eta\alpha)\right] \geq \eta g^{-1}\left[H_b(\alpha)\right]. \tag{18}$$

with equality if and only if $\eta \in \{0,1\}$ or $\alpha = 0$.

Proof. One can see that $g^{-1}\left[H_b(\eta\alpha)\right] = \eta g^{-1}\left[H_b(\alpha)\right]$ if $\eta \in \{0,1\}$ or $\alpha = 0$. In all other cases, we show that

$$g^{-1}\left[H_b(\eta\alpha)\right] > \eta g^{-1}\left[H_b(\alpha)\right]. \tag{19}$$

Let $f(\beta) \triangleq g^{-1}\left[H_b(\beta)\right]$. The Lemma is equivalent to showing that $f(\beta)/\beta$ is a strictly decreasing function in $0 < \beta \leq 1$. Note that since $g(\beta) = H_b(\beta) + 2\left[\log(2) - H_b\left(1/2 + \beta/2\right)\right]$ and $\log(2) > H_b\left(1/2 + \beta/2\right)$ for all $\beta \in (0,1)$, hence $g(\beta) > H_b(\beta)$ for all $0 < \beta < 1$. Since g is one-to-one and increasing, we have $g^{-1}\left[H_b(\beta)\right] < \beta$ for all $0 < \beta < 1$ or $f(\beta) < \beta$ for all $0 < \beta < 1$.

It is not difficult to see that

$$\frac{d}{d\beta}\frac{f(\beta)}{\beta} = \frac{\log\left\{(1-\beta)[1+f(\beta)]\right\}}{\beta^2 \log\left[\frac{1+f(\beta)}{f(\beta)}\right]} \tag{20}$$

and since, using $f(\beta) < \beta$ for all $0 < \beta < 1$, it follows that $(1-\beta)[1+f(\beta)] < 1$ for all $0 < \beta < 1$, hence, $f(\beta)/\beta$ is a strictly decreasing function in $0 < \beta \leq 1$. □

Recall that if the distribution of a random variable X is Binomial, denoted by $\text{Bin}(L, \eta) \in \mathbb{P}$, then $\text{Bin}(L, \eta, k) \triangleq \Pr\{X = k\} = \binom{L}{k}\eta^k(1-\eta)^{L-k}$ if $k \in \{0, 1, ..., L\}$ and is zero otherwise.

Let the two non-zero entries of the probability vector $\boldsymbol{x}^{N,P}$ be at the N-th and P-th position, i.e., $x_N = 1-\alpha, x_P = \alpha$ and let $\boldsymbol{z}^{N,P} = M_\eta(\boldsymbol{x}^{N,P})$.

Lemma 2. *For all $\eta \in [0,1]$, $\alpha \in [0,1]$ and $L \geq 1$, we have*

$$g^{-1}\left[H(\boldsymbol{z}^{N,P})\right] \geq \eta g^{-1}\left[H(\boldsymbol{x}^{N,P})\right]. \tag{21}$$

Proof. By Lemma 1, we have

$$g^{-1}\left[H_b(\eta\alpha)\right] \geq \eta g^{-1}\left[H_b(\alpha)\right]. \tag{22}$$

Note that g is one-one and and strictly increasing, therefore g^{-1} is also strictly increasing. Therefore, it is enough to prove that

$$H(\boldsymbol{z}^{N,P}) \geq H(\boldsymbol{z}^{0,1}). \tag{23}$$

as $H(\boldsymbol{z}^{0,1}) = H_b(\eta\alpha)$ and $H(\boldsymbol{x}^{N,P}) = H_b(\alpha)$. We first show that

$$H(\boldsymbol{z}^{0,P}) \geq H(\boldsymbol{z}^{0,1}). \tag{24}$$

Note that

$$H(\boldsymbol{z}^{0,P}) = f\left[\alpha, (1-\eta)^P\right] + \alpha H\left[\text{Bin}(P, \eta)\right], \tag{25}$$

where

$$f(\alpha,x) = -\left[(1-\alpha)+\alpha x\right]\log\left[(1-\alpha)+\alpha x\right] - (1-x)\alpha\log(\alpha) + x\log(x)\alpha. \tag{26}$$

It is not difficult to show that $f(x)$ is a decreasing function of x. Note that $H\left[\text{Bin}(P,\eta)\right]$ increases with P. Since $H(\boldsymbol{x}^{0,P})$ is a sum of two functions each of which increases with P, (24) follows.

Next, we show that for all $N, P \geq 0$, we have

$$H(\boldsymbol{z}^{N+1,P+1}) \geq H(\boldsymbol{z}^{N,P}). \tag{27}$$

Note first that $\text{Bin}(N+1,\eta) = (1-\eta)\text{Bin}(N,\eta) + \eta\text{Bin}_{+1}(N,\eta)$, where if X has distribution $\text{Bin}_{+1}(N,\eta)$, then $\Pr\{X = k+1\} = \text{Bin}(N,\eta,k)\ \forall\ k$. This implies that

$$\boldsymbol{z}^{N+1,P+1} = (1-\eta)\boldsymbol{z}^{N,P} + \eta\boldsymbol{z}^{N,P}_{+1}, \tag{28}$$

where we define $\boldsymbol{z}^{N,P}_{+1}$ similarly. Using $H(\boldsymbol{z}^{N,P}) = H(\boldsymbol{z}^{N,P}_{+1})$, it is not difficult to show that

$$H(\boldsymbol{z}^{N+1,P+1}) = H(\boldsymbol{z}^{N,P}) + (1-\eta)D\left[\boldsymbol{z}^{N,P}||\boldsymbol{z}^{N+1,P+1}\right] + \eta D\left[\boldsymbol{z}^{N,P}_{+1}||\boldsymbol{z}^{N+1,P+1}\right], \tag{29}$$

where $D(\cdot||\cdot)$ is the relative entropy that is always non-negative and hence, (27) follows.

Assume w.l.o.g. that $P > N$. Applying (24) repeatedly followed by (27), we get

$$H(\boldsymbol{z}^{N,P}) \geq H(\boldsymbol{z}^{0,P-N}) \geq H(\boldsymbol{z}^{0,1}). \tag{30}$$

The result follows. □

4 ρ_A has number states as eigenvectors and ρ_B is the vacuum state

We have observed that the EPnI holds when ρ_A has two non-zero eigenvalues with eigenvectors as the number states and ρ_B is a vacuum state. We now consider the case when ρ_A has number states as the eigenvectors and could have arbitrary number of nonzero eigenvalues and ρ_B is the vacuum state. We derive some necessary and sufficient conditions for this inequality to hold.

We first note that $M_\eta\left[M_{\eta'}(\boldsymbol{x})\right] = M_{\eta\eta'}(\boldsymbol{x})\ \forall\ \eta,\eta' \in [0,1]$ and $\boldsymbol{x} \in \mathbb{P}$. To prove this, let $\boldsymbol{y} = M_{\eta'}(\boldsymbol{x})$, $\boldsymbol{z} = M_\eta(\boldsymbol{y})$ and note that

$$z_i = \sum_{k=i}^{\infty}\binom{k}{i}\eta^i(1-\eta)^{k-i}y_k \tag{31}$$

$$= \sum_{k=i}^{j}\binom{k}{i}\eta^i(1-\eta)^{k-i}\sum_{j=k}^{\infty}\binom{j}{k}(\eta')^k(1-\eta')^{j-k}x_j \tag{32}$$

$$= \sum_{j=i}^{\infty}\binom{j}{i}(\eta\eta')^i x_j \sum_{k-i=0}^{j-i}\binom{j-i}{k-i}(\eta'-\eta\eta')^{k-i}(1-\eta')^{j-k} \tag{33}$$

$$= \sum_{j=i}^{\infty} M_{\eta\eta'}x_j. \tag{34}$$

To simplify the notation, let us define

$$H(\eta, \boldsymbol{x}) \triangleq H(M_\eta x) \tag{35}$$

$$h(\eta, \boldsymbol{x}) \triangleq g^{-1}\left[H(\eta, \boldsymbol{x})\right]. \tag{36}$$

As M_1 is an identity transformation, we sometimes write $H(\boldsymbol{x})$ for $H(1, \boldsymbol{x})$ and $h(\boldsymbol{x})$ for $h(1, \boldsymbol{x})$. Note that $h(1, \boldsymbol{x}) = g^{-1}\left[H(\boldsymbol{x})\right]$ and therefore, (16) can be rephrased as

$$\frac{h(\eta, \boldsymbol{x})}{\eta} \geq h(1, \boldsymbol{x}). \tag{37}$$

It is not difficult to see that if (16) holds, then $h(\eta, \boldsymbol{x})/\eta$ is a decreasing function in η. To see this, let $\eta' \leq \eta$ and $\delta = \eta'/\eta$ where $0 \leq \delta \leq 1$. Then

$$\frac{h(\eta', \boldsymbol{x})}{\eta'} = \frac{h[\delta, M_\eta(\boldsymbol{x})]}{\delta}\frac{1}{\eta} \tag{38}$$

$$\geq \frac{h[1, M_\eta(\boldsymbol{x})]}{\eta} \tag{39}$$

$$= \frac{h(\eta, \boldsymbol{x})}{\eta}. \tag{40}$$

As $h(\eta, \boldsymbol{x})/\eta$ is differentiable, we have

$$\frac{d}{d\eta}\frac{h(\eta, \boldsymbol{x})}{\eta} = \eta\frac{dH(\eta, \boldsymbol{x})}{d\eta} - H(\eta, \boldsymbol{x}) + \log\left[1 + h(\eta, \boldsymbol{x})\right]. \tag{41}$$

Lemma 3. *Let* $M_\eta : [0,1] \times \mathbb{P} \rightarrow \mathbb{P}$ *be the transformation given by* (15). *The following are equivalent:*

$$(i) \qquad h(\eta, \boldsymbol{x}) \geq \eta h(1, \boldsymbol{x}) \qquad \forall\, \boldsymbol{x} \in \mathbb{P}, \forall\, \eta \in (0,1], \tag{42}$$

$$(ii) \qquad \frac{d}{d\eta}\frac{h(\eta, \boldsymbol{x})}{\eta} \leq 0 \qquad \forall\, \boldsymbol{x} \in \mathbb{P}, \forall\, \eta \in (0,1], \tag{43}$$

$$(iii) \qquad \frac{d}{d\eta}\frac{h(\eta, \boldsymbol{x})}{\eta}\Big|_{\eta=1} \leq 0 \qquad \forall\, \boldsymbol{x} \in \mathbb{P}. \tag{44}$$

Proof. It is clear from (40) that (i) and (ii) are equivalent. Furthermore, (ii) implies (iii) since (iii) is a special case of (ii). We prove that (iii) implies (ii). Note that

$$\frac{d}{d\beta}\frac{h[\beta, M_\eta(\boldsymbol{x})]}{\beta}\Big|_{\beta=1} = \frac{d}{d\beta}\frac{h(\eta\beta, \boldsymbol{x})}{\beta}\Big|_{\beta=1} \tag{45}$$

$$= \eta^2\frac{d}{d\theta}\frac{h(\theta, \boldsymbol{x})}{\theta}\Big|_{\theta=\eta}. \tag{46}$$

Now (iii) implies that

$$\frac{d}{d\theta}\frac{h(\theta, \boldsymbol{x})}{\theta}\Big|_{\theta=\eta} \leq 0 \tag{47}$$

and hence, (ii) follows using (46). $\square$

We now state EPnI in (16) in the form of an entropic inequality, i.e., an inequality involving Shannon entropy of discrete probability distributions. By Lemma 3, (16) is equivalent to

$$\eta \frac{dH(\eta,\boldsymbol{x})}{d\eta} - H(\eta,\boldsymbol{x}) + \log\left[1 + h(\eta,\boldsymbol{x})\right] \le 0. \tag{48}$$

The above can be expressed as

$$g\left[\mathrm{e}^{H(\eta,\boldsymbol{x}) - \eta\frac{dH(\eta,\boldsymbol{x})}{d\eta}} - 1\right] \ge H(\eta,\boldsymbol{x}). \tag{49}$$

Note that $g(1/\beta - 1) = H_b(\beta)/\beta \;\forall\; \beta \in [0,1]$ and hence, (16) is equivalent to showing that

$$H(\eta,\boldsymbol{x}) \le \frac{H_b\left[\mathrm{e}^{-H(\eta,\boldsymbol{x}) + \eta\frac{dH(\eta,\boldsymbol{x})}{d\eta}}\right]}{\mathrm{e}^{-H(\eta,\boldsymbol{x}) + \eta\frac{dH(\eta,\boldsymbol{x})}{d\eta}}}. \tag{50}$$

For the two dimensional case with $\eta = 1$, $\boldsymbol{x} = [\alpha, 1-\alpha, 0, ...]$, $\alpha \in [0,1]$, $H(\eta,\boldsymbol{x}) - \eta dH(\eta,\boldsymbol{x})/d\eta = -\log(\alpha)$, $H(\boldsymbol{x}) = H_b(\alpha)$, and substituting this in (50), we get

$$H_b(\alpha) \le \frac{H_b(\alpha)}{\alpha}, \tag{51}$$

which is true. This gives a short proof of (16) for this special case. Evaluating (50) at $\eta = 1$ gives an interesting expression that depends only on the distribution $\boldsymbol{x}$. It is shown in (62) that

$$\Theta(\boldsymbol{x}) \triangleq \frac{dH(\eta,\boldsymbol{x})}{d\eta}\Big|_{\eta=1} = -\sum_{i=1}^{\infty} i x_i \log\left(\frac{x_i}{x_{i-1}}\right), \tag{52}$$

and hence, (50) reduces to

$$H(\boldsymbol{x}) \le \frac{H_b\left[\mathrm{e}^{-H(\boldsymbol{x}) + \Theta(\boldsymbol{x})}\right]}{\mathrm{e}^{-H(\boldsymbol{x}) + \Theta(\boldsymbol{x})}}. \tag{53}$$

The above inequality involves only entropies and another function Θ of the distribution but, to the best of our knowledge, has never been studied before in the literature.

We now show that if (16) is true, then it implies that

$$\eta \frac{dH(\eta,\boldsymbol{x})}{d\eta} < 1, \tag{54}$$

$$\eta \frac{dH(\eta,\boldsymbol{x})}{d\eta} \le H(\eta,\boldsymbol{x}). \tag{55}$$

If (16) holds, then using Lemma 3, we have $H(\eta,\boldsymbol{x}) - \eta dH(\eta,\boldsymbol{x})/d\eta \ge \log\left[1 + h(\eta,\boldsymbol{x})\right]$. As $\log\left[1 + h(\eta,\boldsymbol{x})\right] \ge 0$, we have $H(\eta,\boldsymbol{x}) - \eta dH(\eta,\boldsymbol{x})/d\eta \ge 0$, which proves (55).

Using Lemma 3 again, we have $\eta dH(\eta,\boldsymbol{x})/d\eta - H(\eta,\boldsymbol{x}) + \log\left[1 + h(\eta,\boldsymbol{x})\right] \le 0$. It is enough to prove that $H(\eta,\boldsymbol{x}) - \log\left[1 + h(\eta,\boldsymbol{x})\right] \le 1$, i.e.,

$$1 + g^{-1}\left[H(\eta,\boldsymbol{x})\right] \ge \mathrm{e}^{H(\eta,\boldsymbol{x}) - 1}. \tag{56}$$

We first consider the case when $0 \leq H(\eta, \boldsymbol{x}) \leq 1$. Then $\mathrm{e}^{H(\eta, \boldsymbol{x})-1} \leq 1$. Therefore, $1 + g^{-1}[H(\eta, \boldsymbol{x})] \geq \mathrm{e}^{H(\eta, \boldsymbol{x})-1}$ and (54) holds.

Now consider $H(\eta, \boldsymbol{x}) \geq 1$. Hence, it is enough to prove that $1 + g^{-1}(x) \geq \mathrm{e}^{x-1} \, \forall \, x \geq 1$, or, $x + 1 \geq g(\mathrm{e}^{x} - 1) \, \forall \, x \geq 0$. Simplifying, we can show that this is equivalent to showing that $r(\mathrm{e}^{-x}) \geq 0$, where $r : [0, 1] \to \mathbb{R}$ and

$$r(x) = x + (1 - x) \log(1 - x). \tag{57}$$

Note that $r(0) = 0$ and $dr(x)/dx = -\log(1 - x) \geq 0 \, \forall \, x \in [0, 1]$. Therefore, $r(x) \geq 0 \, \forall \, x \in [0, 1]$ and (54) follows.

(54) and (55) are the necessary conditions for (16) to hold. We now show that they both hold under general conditions.

Lemma 4. *For all $\eta \in [0, 1]$ and $\boldsymbol{x} \in \mathbb{P}$, the following hold:*

$$\eta \frac{dH(\eta, \boldsymbol{x})}{d\eta} < 1, \tag{58}$$

$$\eta \frac{dH(\eta, \boldsymbol{x})}{d\eta} \leq H(\eta, \boldsymbol{x}) \tag{59}$$

with equality if and only if $M_\eta(\boldsymbol{x}) = [1, 0, ...]$.

Proof. Let $\boldsymbol{z} = M_\eta(\boldsymbol{x})$ and using

$$\eta \frac{dz_i}{d\eta} = i z_i - (i + 1) z_{i+1}, \tag{60}$$

we get

$$\begin{aligned} -\eta \frac{dH(\eta, \boldsymbol{x})}{d\eta} &= \eta \sum_{i=0}^{\infty} [1 + \log(z_i)] \frac{dz_i}{d\eta} && (61) \\ &= \sum_{i=1}^{\infty} i z_i \log\left(\frac{z_i}{z_{i-1}}\right) && (62) \\ &\overset{a}{\geq} \sum_{i=1}^{\infty} i z_i \left(1 - \frac{z_{i-1}}{z_i}\right) && (63) \\ &= -1, && (64) \end{aligned}$$

where in a, we have used the inequality that $\log(x) \geq 1 - 1/x$ for all $x \geq 0$ with equality if and only if $x = 1$. If $\boldsymbol{z}$ is such that $z_i \neq 0 \, \forall \, i$, then it is impossible to have an equality in a since equality would imply $z_{i-1} = z_i \, \forall \, i$ and this would imply that $\sum_{i=0}^{\infty} z_i$ is unbounded.

If $\boldsymbol{z}$ has a finite number of nonzero values say $\boldsymbol{z} = [z_0, z_1, ..., z_{L-1}, 0, ...]$, then (64) can be further tightened as

$$\eta \frac{dH(\eta, \boldsymbol{x})}{d\eta} \leq 1 - L z_{L-1}. \tag{65}$$

Hence, (54) holds.

We now prove (59) or equivalently

$$\Theta(\boldsymbol{z}) = -\sum_{i=1}^{\infty} iz_i \log\left(\frac{z_i}{z_{i-1}}\right) \leq H(\boldsymbol{z}). \tag{66}$$

Let us define a sequence of probability distributions $\{\boldsymbol{z}^{(L)}\}$, $L = 0, 1, ...$, where $\boldsymbol{z}^{(L)}$ has length $L+1$ and $\boldsymbol{z}^{(L)} = [(1-z_L)\boldsymbol{z}^{(L-1)}, z_L]$ and $\boldsymbol{z}^{(0)} = [1]$. It is easy to see that the following recurrence relations hold

$$\Theta(\boldsymbol{z}^{(L)}) = (1-z_L)\Theta(\boldsymbol{z}^{(L-1)}) + Lz_L \log\left(\frac{1-z_L}{z_L} z_{L-1}\right) \tag{67}$$

$$H(\boldsymbol{z}^{(L)}) = (1-z_L)H(\boldsymbol{z}^{(L-1)}) + H_b(z_L). \tag{68}$$

Define

$$\Xi(\boldsymbol{z}^{(L)}) \triangleq \Theta(\boldsymbol{z}^{(L)}) - H(\boldsymbol{z}^{(L)}). \tag{69}$$

Using the recurrence relations in (67) and (68), we get

$$\Xi(\boldsymbol{z}^{(L)}) = (1-z_L)\Xi(\boldsymbol{z}^{(L-1)}) + Lz_L \log\left(\frac{1-z_L}{z_L} z_{L-1}\right) - H_b(z_L). \tag{70}$$

We now claim that

$$\Xi(\boldsymbol{z}^{(L)}) \leq L\log(1-z_L). \tag{71}$$

We prove this by induction. It is easy to check that $\Xi(\boldsymbol{z}^{(1)}) = \log(1-z_1)$. Let (71) hold for $L-1$, $L > 1$. Then we have

$$\Xi(\boldsymbol{z}^{(L)}) = (1-z_L)\Xi(\boldsymbol{z}^{(L-1)}) + Lz_L \log\left(\frac{1-z_L}{z_L} z_{L-1}\right) - H_b(z_L) \tag{72}$$

$$\stackrel{a}{\leq} (L-1)(1-z_L)\log(1-z_{L-1}) + (L-1)z_L \log(z_{L-1}) + Lz_L \log\left(\frac{1-z_L}{z_L}\right) - H_b(z_L) \tag{73}$$

$$\stackrel{b}{=} -(L-1)d(z_L, z_{L-1}) + L\log(1-z_L) \tag{74}$$

$$\leq L\log(1-z_L), \tag{75}$$

where in a, we have used the induction hypothesis and the fact that $z_L \log(z_{L-1}) \leq 0$, in b,

$$d(x,y) = x\log\left(\frac{x}{y}\right) + (1-x)\log\left(\frac{1-x}{1-y}\right) \tag{76}$$

is the relative entropy between $[x, 1-x]$ and $[y, 1-y]$ and is always nonnegative. (59) now follows from (71) since $\log(1-z_L) \leq 0$. The equality condition follows straightforwardly. □

It is not difficult to see that the sufficient condition for (16) to hold is that $dH(\eta, \boldsymbol{x})/d\eta \leq 0$. This condition is, of course, not true for many distributions such as a distribution whose sequence of entries are non-increasing. Suppose $\boldsymbol{z} = M_\eta(\boldsymbol{x})$ has some zero entries in its interior, i.e.,

$z_i = 0$ and $z_{i+1} \neq 0$ for some i. Then one can easily check that $dH(\eta, \boldsymbol{x})/d\eta = -\infty$ and (16) holds. It also follows from (65) that if, for distributions with finite non-zero entries of the form $\boldsymbol{z} = [z_0, z_1, ..., z_{L-1}, 0, ...]$ and $z_{L-1} \geq 1/L$, then (16) holds.

We now show that (16) holds if $H(\boldsymbol{x})$ is sufficiently large.

Lemma 5. *For a given $\eta \in (0,1)$, $\boldsymbol{x} \in \mathbb{P}$,* (16) *holds if $H(\boldsymbol{x})$ is large enough.*

Proof. Using (49), we need to show that

$$g\left[\mathrm{e}^{H(\eta,\boldsymbol{x})-\eta dH(\eta,\boldsymbol{x})/d\eta} - 1\right] \geq H(\eta, \boldsymbol{x}). \tag{77}$$

We have

$$g\left[\mathrm{e}^{H(\eta,\boldsymbol{x})-\eta dH(\eta,\boldsymbol{x})/d\eta} - 1\right] \overset{a}{>} H(\eta,\boldsymbol{x}) + \delta - \mathrm{e}^{-H(\eta,\boldsymbol{x})+\eta dH(\eta,\boldsymbol{x})/d\eta} \tag{78}$$

$$\overset{b}{>} H(\eta,\boldsymbol{x}) + \delta - \mathrm{e}^{-H(\eta,\boldsymbol{x})+1} \tag{79}$$

$$\geq H(\eta,\boldsymbol{x}), \tag{80}$$

where in a, we use the inequality that $g(\mathrm{e}^x - 1) \geq x + 1 - \mathrm{e}^{-x}$ and we use Lemma 4 to get $\eta dH(\eta,\boldsymbol{x})/d\eta < 1 - \delta$ for some $\delta > 0$, in b, we use $\eta dH(\eta,\boldsymbol{x})/d\eta < 1$ and the last inequality would hold if $H(\eta,\boldsymbol{x}) \geq 1 - \log(\delta)$ or if $H(\eta,\boldsymbol{x})$ is large enough.

We now show that if $H(\boldsymbol{x})$ is large, then so is $H(\eta,\boldsymbol{x})$ for $\eta \in (0,1)$. Define

$$q(\eta,\boldsymbol{x}) \triangleq \frac{H(\eta,\boldsymbol{x})}{\eta}. \tag{81}$$

Differentiating w.r.t. η, we get using (59),

$$\frac{dq(\eta,\boldsymbol{x})}{d\eta} = \frac{1}{\eta^2}\left[\eta\frac{dH(\eta,\boldsymbol{x})}{d\eta} - H(\eta,\boldsymbol{x})\right] \tag{82}$$

$$\leq 0. \tag{83}$$

Hence, $q(\eta,\boldsymbol{x})$ is a decreasing function of η and $H(\eta,\boldsymbol{x}) \geq \eta H(\boldsymbol{x})$. Similarly, using (58), we get

$$\int_\eta^1 dH(\beta,\boldsymbol{x}) < \int_\eta^1 \frac{d\beta}{\beta} \tag{84}$$

$$H(\eta,\boldsymbol{x}) > H(\boldsymbol{x}) + \log(\eta). \tag{85}$$

Hence,

$$H(\eta,\boldsymbol{x}) \geq \max\left\{\eta H(\boldsymbol{x}), H(\boldsymbol{x}) + \log(\eta)\right\}. \tag{86}$$

This shows that if $H(\boldsymbol{x})$ is large, then so is $H(\eta,\boldsymbol{x})$ and hence, (16) would hold for any $\eta \in (0,1]$ for large $H(\boldsymbol{x})$. □

5 Discussion

It is, of course, of great interest to see if these results could be generalized for the cases where ρ_A and ρ_B do not have the special structure such as the eigenvectors being the number states etc. It would seem that our results may extend over to cover some of these cases if the following is established. Suppose there exists an $\boldsymbol{x} \in \mathbb{P}$ such that

$$\frac{d}{d\eta}\frac{h[\eta, M_\beta(\boldsymbol{x})]}{\eta}\Big|_{\eta=1} < 0 \;\forall\, \beta \in (0,1]. \tag{87}$$

Then, it follows from (46) that

$$\frac{d}{d\eta}\frac{h(\eta, \boldsymbol{x})}{\eta}\Big|_{\eta=\beta} < 0 \;\forall\, \beta \in (0,1]. \tag{88}$$

This would then imply that $h(\beta, \boldsymbol{x})$ is a strictly decreasing function of $\beta \in (0,1]$ and hence, (16) holds with strict inequality.

An example of such a $\boldsymbol{x}$ is $\boldsymbol{x} = [\alpha, 1-\alpha, 0, ...]$, $\alpha \neq 1$, and $M_\beta(\boldsymbol{x}) = [1-(1-\alpha)\beta, (1-\alpha)\beta, 0, ...]$, and (16) is strict using (51).

For finite n and any state σ defined on the number states as

$$\sigma = \sum_{i,j=0}^{n} \xi_{i,j} \left|i\right\rangle \left\langle j\right|, \tag{89}$$

we define a function

$$f(n,\sigma) = \sum_{i,j=0}^{n} \xi_{i,j} \left|e_i^n\right\rangle \left\langle e_j^n\right|, \tag{90}$$

where $\{|e_i^n\rangle\}$ is the standard basis for the Hilbert space of dimension $n+1$, i.e., $\langle e_i| = [\overbrace{0,...,0}^{i}, 1, \overbrace{0,...,0}^{n-i}]$, $i = 0, 1, ..., n$. It follows that $S(\sigma) = S[f(n,\sigma)]$.

Now consider the input states such that

$$\rho_A = \sum_{i,j=0}^{n_A} \lambda_{i,j} \left|i\right\rangle_A \left\langle j\right|_A, \tag{91}$$

$$\rho_B = \sum_{i,j=0}^{n_B} \gamma_{i,j} \left|i\right\rangle_B \left\langle j\right|_B, \tag{92}$$

$$\hat{\rho}_A = \alpha \left|0\right\rangle_A \left\langle 0\right|_A + (1-\alpha) \left|1\right\rangle_A \left\langle 1\right|_A \tag{93}$$

$$\hat{\rho}_B = \left|0\right\rangle_B \left\langle 0\right|_B, \tag{94}$$

where n_A, n_B are finite and $||f(n_A,\rho_A) - f(n_A,\hat{\rho}_A)||_{\mathrm{tr}} < \delta$ and $||f(n_B,\rho_B) - f(n_B,\hat{\rho}_B)||_{\mathrm{tr}} < \delta$.

It is not difficult to see that under the action of f, the output ρ_C of beam splitter with ρ_A and ρ_B as inputs is close to the output $\hat{\rho}_C$ with $\hat{\rho}_A$ and $\hat{\rho}_B$ as inputs, i.e., $||f(n_A+n_B,\rho_C) -$

$f(n_A + n_B, \hat{\rho}_C)||_{\mathrm{tr}} < \epsilon$, where we could make ϵ as small as possible by choosing δ small. Using Fannes' inequality [9, 10], this would result in a small deviation in the von Neumann entropies of ρ_A, ρ_B and ρ_C as compared to $\hat{\rho}_A$, $\hat{\rho}_B$ and $\hat{\rho}_C$ respectively that can be absorbed while still preserving the inequality since the inequality is strict.

References

[1] S. Guha, B. I. Erkemen, and J. H. Shapiro, "The Entropy Photon-Number Inequality and its Consequences," *Open Problems Session, ITA, UCSD*, 2008.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–655, July and Oct. 1948.

[3] A. J. Stam, "Some inequalities satisfied by the quantities of information of fisher and shannon," *Inf. Contr.*, vol. 2, pp. 101–112, June 1959.

[4] N. M. Blachman, "The convolution inequality for entropy powers," *IEEE Trans. Inf. Theory*, vol. 11, pp. 267–271, Apr. 1965.

[5] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, "Classical capacity of the lossy bosonic channel: The exact solution," *Phys. Rev. Lett.*, vol. 92, p. 027902, Jan 2004.

[6] S. Guha, J. H. Shapiro, and B. I. Erkemen, "Classical capacity of bosonic broadcast communication and a minimum output entropy conjecture," *Phys. Rev. A*, vol. 76, no. 032303, 2007.

[7] ——, "Capacity of the bosonic wiretap channel and the Entropy Photon-Number Inequality," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Toronto, Canada, July 2008.

[8] C. Gerry and P. Knight, *Introductory Quantum Optics*. Cambridge University Press, 2004.

[9] M. Fannes, "A continuity property of the entropy density for spin lattice systems," *Commun. Math. Phys.*, vol. 31, pp. 291–294, 1973.

[10] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quanum Information*. Cambridge U.P., 2000.