On some special cases of the Entropy Photon-Number Inequality

Smarajit Das, Naresh Sharma and Siddharth Muthukrishnan

Tata Institute of Fundamental Research

December 5, 2011

One of the aims of information theory is to determine what is the capacity of a noisy communication channel.

Capacity is defined as the highest communication rate that can be sent across a communication channel reliably [Shannon, 1948].

Shannon theory has been applied to the quantum case and one can define the capacity as the upper limit on the qubits that can be transferred across a quantum channel reliably.

Since Shannon's work, capacity of various channels has been determined and for many channels (both classical and quantum) capacity is still unknown.

A zoo of capacities exists under different assumptions such as with or without shared entanglement between Alice (sender) and Bob (receiver), for sending classical and/or quantum data, point to point or network scenarios etc.

Some examples of the quantum channels whose capacities are not known

1. Bosonic thermal noise channel [Giovannetti *et al*, 2004]

2. Bosonic broadcast channel [Guha *et al*, 2007]

If the entropy photon-number inequality (EPnI) is true, then it implies certain conjectures that would establish the capacity of above two channels.

What is EPnI and what conjectures it implies?

The Entropy Photon Number Inequality (EPnI) was conjectured by Guha, Shapiro and Erkmen (2008).

EPnI has a classical analogue called the 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 for $\eta \in [0, 1]$,

$$\mathrm{e}^{2h(\sqrt{\eta}X+\sqrt{1-\eta}Y)} \geqslant \eta\, \mathrm{e}^{2h(X)} + (1-\eta)\, \mathrm{e}^{2h(Y)}.$$

First stated by Shannon [1948] and the proof was given by Stam [1959] and Blachman [1965].

$\mathrm{e}^{2h(X)}$ is the power (variance) of a Gaussian random variable having the same entropy as X.

For bosonic channels, we need to look at the creation and annihilation operators for bosons.

Same as the ladder operators of the quantum harmonic oscillator.

a is the annihilation operator given by

$$a = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The creation operator is $a^\dagger$.

The number operator is $N = a^\dagger a$.

$[A, B] := AB - BA$.

$[a, a^\dagger] = \mathbf{1}$, $[N, a^\dagger] = a^\dagger$, $[N, a] = -a$.

Fock/number states $\{|0\rangle, |1\rangle, |2\rangle, ...\}$ form an orthonormal basis of the underlying Hilbert space and

$$\begin{aligned} N|n\rangle &= n|n\rangle, \\ a^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle, \\ a|n\rangle &= \sqrt{n}|n-1\rangle, \qquad a|0\rangle = \mathbf{0}, \\ |n\rangle &= \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle. \end{aligned}$$

A thermal state with mean photon-number μ is given by

$$\rho_T = \sum_{i=0}^{\infty} \frac{\mu^i}{(1+\mu)^{i+1}} |i\rangle\langle i|.$$

Consider a beam-splitter with two inputs and two outputs.

For simplicity, we shall assume that each input has one boson each and let the annihilation operators for inputs 1 and 2 be given by a and b respectively.

The joint state associated with a and b is the product state $\rho_{AB} = \rho_A \otimes \rho_B$.

Let the annihilation operators associated with the outputs be c and d and

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \sqrt{\eta} & \sqrt{1-\eta} \\ -\sqrt{1-\eta} & \sqrt{\eta} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

η, $1-\eta$ are the transmissivity and reflectivity of the beam splitter.

Calculation of the state at the output is outlined as follows. $|i\rangle_Z$ are the Fock number states for $Z \in \{A, B, C, D\}$.

Note

$$|i\rangle_A|j\rangle_B = \frac{(a^\dagger)^i}{\sqrt{i!}}\frac{(b^\dagger)^j}{\sqrt{j!}}|0\rangle_A|0\rangle_B,$$

$$|i\rangle_A|j\rangle_B \xrightarrow{\text{B.S.}} \frac{(\sqrt{\eta}c^\dagger + \sqrt{1-\eta}d^\dagger)^i}{\sqrt{i!}}\frac{(\sqrt{1-\eta}c^\dagger - \sqrt{\eta}d^\dagger)^j}{\sqrt{j!}}|0\rangle_C|0\rangle_D.$$

B.S. indicates the action of the beam splitter.

$c^\dagger$ and $d^\dagger$ commute. Using the binomial expansion, we get

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

# of summations is 4 when looking at the action of the beam splitter on the states of the form $|i\rangle_A|j\rangle_B\langle\acute{i}|_A\langle\acute{i}|_B$.

If ρ_A and ρ_B have eigenvectors as the Fock number states, expression for ρ_C involves 5 summations (9 otherwise) - one summation disappears due to orthogonality of the number states.

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

The EPnI is now stated as

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

$S(\rho) = -\operatorname{Tr}(\rho \log \rho)$ is the von Neumann entropy.

$g(x) = (x+1)\log(x+1) - x\log(x)$ is the von Neumann entropy of the thermal state with mean photon-number x.

$g^{-1}[S(\rho)]$ is the mean photon-number of the thermal state having the same entropy as ρ.

EPnI implies minimum output entropy conjectures that can be used to find the capacity of certain channels.

Minimum Output Entropy Conjecture 1: If ρ_A is a pure state ($\rho_A = |\psi\rangle\langle\psi|$) and ρ_B is a thermal state, then $S(\rho_C)$ is minimized by choosing $\rho_A = |0\rangle\langle 0|$ (the vacuum state).

Minimum Output Entropy Conjecture 2: If ρ_A is a pure state ($\rho_A = |\psi\rangle\langle\psi|$) and $S(\rho_B) = g(K)$ for some $K \geqslant 0$, then $S(\rho_C)$ is minimized by choosing ρ_B to be a thermal state with mean photon-number K.

These conjectures are corollaries of EPnI and are yet to be proved independently.

We examine if EPnI holds for some special cases.

$\rho_B = |0\rangle\langle 0|$ is a vacuum state.

$\mathbb{P}$ - set of all probability distributions defined on nonnegative integers.

Let the eigenvalues of ρ_A be given by the probability vector $\boldsymbol{x} \in \mathbb{P}$.

$$\rho_A = \sum_{i=0}^{\infty} x_i |i\rangle_A \langle i|_A,$$

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

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

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

EPnI reduces to

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

Holds trivially for $\eta = 0, 1$.

We concentrate on this special inequality for most part.

One might think that with > 60 years of information theory, one would just pick an inequality from an information theory book whose corollary (if not a corollary of a corollary) would be the above inequality.

$$M_\eta(\mathbf{x}) = \begin{pmatrix} 1 & (1-\eta) & (1-\eta)^2 & (1-\eta)^3 & \cdots \\ 0 & \eta & 2\eta(1-\eta) & 3\eta(1-\eta)^2 & \cdots \\ 0 & 0 & \eta^2 & 3\eta^2(1-\eta) & \cdots \\ 0 & 0 & 0 & \eta^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ \vdots \end{pmatrix}$$

The eigenvalues of ρ_A given by the probability vector $\boldsymbol{x} = [1-\alpha, \alpha, 0, 0, ...], \alpha \in [0,1]$.

For $p \in [0,1]$, binary entropy is $H_b(p) := -p\log(p) - (1-p)\log(1-p)$.

EPnI is now equivalent to showing that for all $\eta \in [0,1]$ and $\alpha \in [0,1]$, we have

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

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

This holds.

Let ρ_A have number states as eigenvectors and the number is arbitrary.

Can show that $M_\eta \left[M_{\eta'}(\boldsymbol{x})\right] = M_{\eta\eta'}(\boldsymbol{x}) \; \forall \; \eta, \eta' \in [0,1]$ and $\boldsymbol{x} \in \mathbb{P}$.

Define

$$H(\eta, \boldsymbol{x}) := H(M_\eta x)$$
$$h(\eta, \boldsymbol{x}) := g^{-1}\left[H(\eta, \boldsymbol{x})\right].$$

Then EPnI can be rewritten as $h(\eta, \boldsymbol{x}) \geqslant \eta h(1, \boldsymbol{x})$.

Can be shown to be equivalent to $h(\eta, \boldsymbol{x})/\eta$ is a decreasing function in $\eta \in (0,1]$ or

$$\frac{d}{d\eta}\frac{h(\eta, \boldsymbol{x})}{\eta} \leqslant 0.$$

By using the above multiplicative property, we could just look at $\eta = 1$ to show that EPnI is equivalent to

$$\frac{d}{d\eta}\frac{h(\eta, \boldsymbol{x})}{\eta}\Big|_{\eta=1} \leqslant 0.$$

$$\frac{d}{d\eta}\frac{h(\eta, \boldsymbol{x})}{\eta} = \eta\frac{dH(\eta, \boldsymbol{x})}{d\eta} - H(\eta, \boldsymbol{x}) + \log\left[1 + g^{-1}\left[H(\eta, \boldsymbol{x})\right]\right] \leqslant 0.$$

Rewrite as

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

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

Easy to check that for the 2-dim case with $\eta = 1$, $\boldsymbol{x} = [\alpha, 1-\alpha, 0, ...]$, $\alpha \in [0, 1]$, the above reduces to

$$H_b(\alpha) \leqslant \frac{H_b(\alpha)}{\alpha}.$$

Some more properties

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

$$\eta \frac{dH(\eta, \boldsymbol{x})}{d\eta} \leqslant H(\eta, \boldsymbol{x}).$$

Recall that EPnI would hold if

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

Using the fact that for all $z\geqslant 0$,

$$g\left(\mathrm{e}^{z}-1\right)\geqslant z,$$

we arrive at the following sufficient condition for the EPnI to hold

$$\frac{dH(\eta,\boldsymbol{x})}{d\eta}\leqslant 0.$$

Note

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

Holds surely for the following distributions

1. The entries of $\mathbf{x}$ are non-increasing, i.e., $x_{i+1} \geqslant x_i \ \forall \ i \geqslant 0$.
2. $\mathbf{x}$ has some zero entries in its interior, i.e., $x_i = 0$ and $x_{i+1} \neq 0$ for some i.
3. If $\mathbf{x}$ has finite non-zero entries of the form $\mathbf{x} = [x_0, x_1, ..., x_{L-1}, 0, ...]$ and $x_{L-1} \geqslant 1/L$.

EPnI holds if $H(\boldsymbol{x})$ is sufficiently large.

Outline of the proof:

Can show that if $H(\boldsymbol{x})$ is large, then so is $H(\eta, \boldsymbol{x})$.

$g(\mathrm{e}^x - 1) \geqslant x + 1 - \mathrm{e}^{-x}$

$\exists$ a $\delta > 0$ such that $\eta dH(\eta, \boldsymbol{x})/d\eta < 1 - \delta$.

$H(\eta, \boldsymbol{x}) \geqslant 1 - \log(\delta)$ if $H(\eta, \boldsymbol{x})$ is large enough.

Using the above claims, we get

$$\begin{aligned} g\left[\mathrm{e}^{H(\eta,\boldsymbol{x})-\eta dH(\eta,\boldsymbol{x})/d\eta}-1\right] &> H(\eta,\boldsymbol{x})+\delta-\mathrm{e}^{-H(\eta,\boldsymbol{x})+\eta dH(\eta,\boldsymbol{x})/d\eta} \\ &> H(\eta,\boldsymbol{x})+\delta-\mathrm{e}^{-H(\eta,\boldsymbol{x})+1} \\ &\geqslant H(\eta,\boldsymbol{x}), \end{aligned}$$

which proves EPnI.

Thus far we have shown that EPnI holds only for special cases.

Say EPnI holds for ρ_A and ρ_B with strict inequality.

Suppose we work with states that are close (say in trace distance) to ρ_A and ρ_B. Would EPnI hold for these states as well?

We can invoke the continuity of entropy (Fannes' inequality) and that beam splitter output is a continuous function of the inputs to claim that it would hold for these states as well.

Showed that EPnI holds for special cases.

Give conditions which if satisfied would extend these results.

Things are so messed up (9 summations in full generality) that perhaps a different approach should be tried to attack the problem in a general setting.