A generalization of the Entropy Power Inequality to Bosonic Quantum Systems

G. De Palma,[1,2] A. Mari,[1] and V. Giovannetti[1]

[1] *NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, I-56127 Pisa, Italy.*

[2] *INFN, Pisa, Italy*

In most communication schemes information is transmitted via travelling modes of electromagnetic radiation. These modes are unavoidably subject to environmental noise along any physical transmission medium and the quality of the communication channel strongly depends on the minimum noise achievable at the output. For classical signals such noise can be rigorously quantified in terms of the associated Shannon entropy and it is subject to a fundamental lower bound called entropy power inequality. Electromagnetic fields are however quantum mechanical systems and then, especially in low intensity signals, the quantum nature of the information carrier cannot be neglected and many important results derived within classical information theory require non-trivial extensions to the quantum regime. Here we prove one possible generalization of the Entropy Power Inequality to quantum bosonic systems. The impact of this inequality in quantum information theory is potentially large and some relevant implications are considered in this work.

I.   INTRODUCTION

In standard communication schemes, even if based on a digital encoding, the signals which are physically transmitted are intrinsically analogical in the sense that they can assume a continuous set of values. For example, the usual paradigm is the transmission of information via amplitude and phase modulation of an electromagnetic field. In general, a continuous signal with k components can be modeled by a random variable X with values in $\mathbb{R}^k$ associated with a probability measure $d\mu(x) = p(x)d^k x$ on $\mathbb{R}^k$. For example, a single mode of electromagnetic radiation is determined by a complex amplitude and therefore it can be classically described by a random variable X with $k = 2$ real components. The Shannon differential entropy[1,2] of a general random variable X is defined as

$$H(X) = -\int_{\mathbb{R}^k} p(x) \ln p(x)\, d^k x \,, \quad x \in \mathbb{R}^k \,, \tag{1}$$

and plays a fundamental role in information theory. Indeed depending on the context $H(X)$ quantifies the noise affecting the signal or, alternatively, the amount of information potentially encoded in the variable X.

Now, assume to *mix* two random variables A and B and to get the new variable $C = \sqrt{\lambda}\, A + \sqrt{1-\lambda}\, B$ with $\lambda \in [0,1]$ (see Fig. 1). For example this is exactly the situation in which two optical signals are physically mixed via a beam–splitter of transmissivity λ. What can be said about the entropy of the output variable C? It can be shown that, if the inputs A and B are independent, the following *Entropy Power Inequality* (EPI) holds[3,4]

$$e^{2H(C)/k} \geq \lambda\, e^{2H(A)/k} + (1 - \lambda)\, e^{2H(B)/k} \,, \tag{2}$$

stating that for fixed $H(A)$, $H(B)$, the output entropy $H(C)$ is minimized taking A and B Gaussian with proportional covariance matrices. This is basically a lower bound on $H(C)$ and the name *entropy power* is motivated by the fact that if $p(x)$ is a product of k equal isotropic Gaussians one has $\frac{1}{2\pi e} e^{2H(X)/k} = \sigma^2$, where σ^2 is the variance of each Gaussian which is usually identified with the energy or *power* of the signal[1]. In the context of (classical) probability theory, several equivalent reformulations[2] and generalizations[5–7] of Eq. (2) have been proposed, whose proofs have recently renewed the interest in the field. As a matter of fact, these inequalities play a fundamental role in classical information theory, by providing computable bounds for the information capacities of various models of noisy channels[1,8,9].
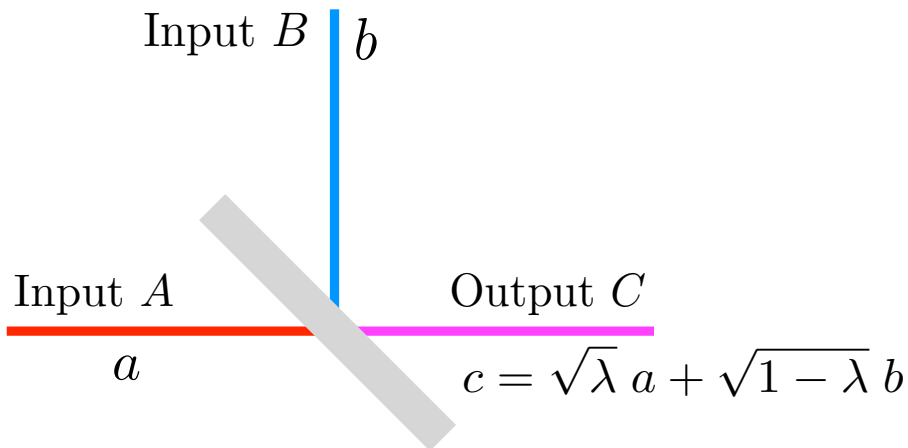
FIG. 1. Graphical representation of the coherent mixing of the two inputs A and B. For the quantum mechanical analogue the two input signals correspond to electromagnetic modes which are coherently mixed at a beam–splitter of transmissivity λ. The entropy of the output signal is lower bounded by a function of the input entropies via the quantum entropy power inequality defined in Eq. (5).

The need for a quantum version of the EPI has arisen in the attempt of solving some fundamental problems in quantum communication theory. In particular the EPI has come into play when it has been realized that a suitable generalization to the quantum setting, called *Entropy Photon number Inequality* (EPnI)[10,11], would directly imply the solution of several optimization problems, including the determination of the classical capacity of Gaussian channels and of the capacity region of the bosonic broadcast channel[12,13]. Up to now the EPnI is still unproved and, while the classical capacity has been recently computed[14,15] by proving the bosonic minimum output entropy conjecture[16], the exact capacity region of the broadcast channel remains undetermined. In 2012 another quantum generalization of the EPI has been proposed, called *quantum Entropy Power Inequality* (qEPI)[17,18], together with its proof valid only for the $50:50$ beam–splitter corresponding to the case $\lambda = \frac{1}{2}$. The contribution of this paper is to show the validity of this inequality for any beam–splitter and to extend it also to the quantum amplifier.

The qEPI proved in this work, while directly giving tight bounds on several entropic quantities, also constitutes a potentially powerful tool which could be used in quantum

information theory in the same spirit in which the classical EPI was instrumental in deriving important classical results like: a bound to the capacity of non–Gaussian channels[1], the convergence of the central limit theorem[19], the secrecy capacity of the Gaussian wire–tap channel[9], the capacity region of broadcast channels[8], *etc.*. In this work we consider some of the direct consequences of the qEPI and we hope to stimulate the research of other important implications in the field.

II. THE QUANTUM ENTROPY POWER INEQUALITY AND ITS PROOF

In order to define the quantum mechanical analogue of Eq. (2), we follow the reasoning line of Ref.s 10, 11, 17, and 18 where the classical random variable X is replaced by a collection of independent bosonic modes. Specifically consider n optical modes described by $a_1, a_2, ..., a_n$ annihilation operators obeying the bosonic commutation rules $\left[a_i, a_j^\dagger\right] = \delta_{ij}$[20,21]. This system represents the quantum analogue of the classical random variable A. We observe that, since the phase space of each mode is 2-dimensional, the total number of phase space variables is $2n$ and this should be identified with the number k appearing in the classical EPI (2). A similar collection of bosonic modes $b_1, b_2, ..., b_n$ will play the role of system B. The natural way of mixing the two signals is via a beam–splitter of transmissivity λ[22], which in the quantum optics formalism is represented by the unitary operation $U = e^{\arctan\sqrt{\frac{1-\lambda}{\lambda}}\sum_j\left(a_j^\dagger b_j - a_j b_j^\dagger\right)}$. This produces n output modes with bosonic operators

$$c_j = \sqrt{\lambda}\, a_j + \sqrt{1-\lambda}\, b_j, \quad j = 1, 2, ..., n . \tag{3}$$

In the Schrödinger picture the above transformation corresponds to a quantum channel[23] mapping the input state ρ_{AB} to the output state

$$\rho_C = \mathcal{E}(\rho_{AB}) = \text{Tr}_B\left[U\rho_{AB}U^\dagger\right] , \tag{4}$$

where the partial trace Tr_B stems for the fact that we discard one of the two output ports of the beam–splitter. We consider the case of independent inputs A, B, with a factorized density matrix $\rho_{AB} = \rho_A \otimes \rho_B$. The *quantum Entropy Power Inequality* (qEPI) reads then

$$e^{S(\rho_C)/n} \geq \lambda\, e^{S(\rho_A)/n} + (1 - \lambda)\, e^{S(\rho_B)/n} , \tag{5}$$

where the classical Shannon entropy has been replaced by the quantum von Neumann entropy $S(\rho) = -\text{Tr}\left[\rho\ln\rho\right]$. Unlike the classical case, the qEPI is *not* saturated by Gaussian states with proportional covariance matrices, unless they have the same entropy. The

qEPI (5) was conjectured in Ref. 18 where it was shown to hold only for the special case of $\lambda = \frac{1}{2}$. In this work we prove that inequality (5) is indeed valid for every λ. Moreover we extend the qEPI to the case in which the two input states are mixed via a *quantum amplifier, i.e.* when the unitary U is replaced by the two-mode squeezing[22] operation $U' = e^{\text{arctanh} \sqrt{\frac{\kappa-1}{\kappa}} \sum_j \left(a_j^\dagger b_j^\dagger - a_j b_j\right)}$ with $\kappa \in [1, \infty]$. In this case the modes a_j are amplified and the modes b_j are phase-conjugated. In the Heisenberg picture we get

$$c_j = \sqrt{\kappa}\, a_j + \sqrt{\kappa - 1}\, b_j^\dagger, \quad j = 1, 2, ..., n\ , \tag{6}$$

and the amplifier version of the qEPI becomes

$$e^{S(\rho_C)/n} \geq \kappa\, e^{S(\rho_A)/n} + (\kappa - 1)\, e^{S(\rho_B)/n}\ . \tag{7}$$

In the rest of the paper we are going to prove the validity of both inequalities (5) and (7) and to show some of their direct implications.

A. Properties of quantum Fisher information

Almost all classical proofs[6] of the EPI are based on two properties of the Fisher information: the *Fisher information inequality* (or *Stam inequality*[4,24]) and the *de Bruijn identity*[3]. Here we follow the approach of 17 in order to generalize such properties to quantum systems. Given a smooth family of states $\theta \mapsto \rho^{(\theta)}$ the associated quantum Fisher information can be defined in terms of the relative entropy:

$$J\left(\rho^{(\theta)}; \theta\right)\big|_{\theta=0} \equiv \frac{d^2}{d\theta^2} S\left(\rho^{(0)} \,\|\, \rho^{(\theta)}\right)\bigg|_{\theta=0}\ , \tag{8}$$

where $S(\rho_1 \| \rho_2) \equiv \text{Tr}\left[\rho_1 \left(\ln \rho_1 - \ln \rho_2\right)\right]$. Since the relative entropy is non-negative and vanishes for $\theta = 0$, we necessarily have $J\left(\rho^{(\theta)}; \theta\right) \geq 0$ and from the definition it is clear that $J\left(\rho^{(c\theta)}; \theta\right) = c^2 J\left(\rho^{(\theta)}; \theta\right)$. Moreover, from the data processing inequality for the relative entropy its counterpart for the quantum Fisher information follows: for every quantum channel $\mathcal{E}$, one has $J\left(\mathcal{E}\left(\rho^{(\theta)}\right); \theta\right) \leq J\left(\rho^{(\theta)}; \theta\right)$[17]. For our purposes, the relevant cases are when θ is associated with translations along the phase space axes, *i.e.* $\rho^{(q,P_j)} = e^{iqP_j} \rho\, e^{-iqP_j}$ and $\rho^{(p,Q_j)} = e^{-ipQ_j} \rho\, e^{ipQ_j}$, where as usual $Q_j \equiv \left(a_j + a_j^\dagger\right)/\sqrt{2}$, $P_j \equiv i\left(a_j^\dagger - a_j\right)/\sqrt{2}$. In this situation one can generalize important results form classical information theory. In particular if two input states are mixed via a beam–splitter or via a quantum amplifier as

described in Eq. (4), one can derive from the data processing inequality the quantum version of the *Stam inequality*:

$$\frac{1}{J_C} \geq \frac{\lambda_A}{J_A} + \frac{\lambda_B}{J_B} \ , \tag{9}$$

where $J = \sum_j J\left(\rho^{(p,Q_j)}; p\right) + J\left(\rho^{(q,P_j)}; q\right)$ and

$$\lambda_A \equiv \lambda \qquad \lambda_B \equiv 1 - \lambda \qquad \text{(beam–splitter)} \tag{10}$$

$$\lambda_A \equiv \kappa \qquad \lambda_B \equiv \kappa - 1 \qquad \text{(amplifier) .} \tag{11}$$

The proof of (9) for the beam–splitter in the special case $\lambda = \frac{1}{2}$ was given in Ref. 17. The key point of this paper is the generalization of this proof to any beam–splitter and amplifier, crucial for the derivation of the qEPIs (5) and (7). In 17, (9) is derived from the inequality

$$w_C^2 J_C \leq w_A^2 J_A + w_B^2 J_B \qquad \forall \ w_A, \ w_B \in \mathbb{R} \ , \tag{12}$$

$$w_C = \sqrt{\lambda_A} \ w_A + \sqrt{\lambda_B} \ w_B \ , \tag{13}$$

proven for any beam–splitter (see Methods for the proof in the amplifier case). Our main idea is to choose w_A and w_B in order to get from (12) the strongest possible inequality. For this purpose, we can rewrite w_C^2 as

$$w_C^2 = \left(\sqrt{\frac{\lambda_A}{J_A}} \ w_A \sqrt{J_A} + \sqrt{\frac{\lambda_B}{J_B}} \ w_B \sqrt{J_B} \right)^2 \leq$$

$$\leq \left(\frac{\lambda_A}{J_A} + \frac{\lambda_B}{J_B} \right) \left(w_A^2 J_A + w_B^2 J_B \right) \ , \tag{14}$$

where we have used the Cauchy–Schwarz inequality. Equality holds iff

$$w_A = k \ \frac{\sqrt{\lambda_A}}{J_A} \ , \qquad w_B = k \ \frac{\sqrt{\lambda_B}}{J_B} \ , \qquad k \in \mathbb{R} \ , \tag{15}$$

and with this choice (12) becomes exactly the generalized Stam inequality (9).

Another important and useful property is the quantum analogue of the *de Bruijn identity* which relates the Fisher information to the the entropy flow under additive Gaussian noise,

$$J \equiv \sum_j J\left(\rho(t)^{(q,P_j)}; q\right) + J\left(\rho(t)^{(p,Q_j)}; p\right) = 4\frac{d}{dt}S(\rho(t)) \ , \tag{16}$$

where $\rho(t) = e^{\mathcal{L}t}\rho(0)$ and

$$\mathcal{L}(\rho) \equiv -\frac{1}{4} \sum_{j=1}^{n} \left([Q_j, [Q_j, \rho]] + [P_j, [P_j, \rho]]\right) \ . \tag{17}$$

The proof, repeated in the Methods, simply follows from the definition of the ensembles $\rho^{(p,Q_j)}$ and $\rho^{(q,P_j)}$[17].

B.   Proof of quantum Entropy Power Inequality

The argument is similar to the one used in the derivation of the classical EPI. This technique, which is based on the addition of white Gaussian noise in the system, was extended to the quantum domain in Ref. 17 in order to prove the qEPI for the special case of $\lambda = \frac{1}{2}$. Here we use the properties (9) and (16) of the quantum Fisher information and we show that the qEPI is valid for all $\lambda \in [0, 1]$ (beam–splitter) and for all $\kappa \geq 1$ (amplifier).

The key idea borrowed from the classical proof is to notice that, for highly entropic thermal states, inequalities (5) and (7) are almost saturated. Then if we evolve the inputs adding classical Gaussian noise, (5) and (7) will asymptotically hold in the infinite time limit, and we just need to prove that the added noise has not improved the inequalities. This can be achieved in the quantum setting by the application of the Gaussian additive noise channel

$$\rho(t) \equiv e^{t\mathcal{L}}\rho \,, \tag{18}$$

where the Liouvillian operator $\mathcal{L}$ is the one defined in Eq. (17). We need an asymptotic estimate for the entropy of $\rho(t)$ as $t \to \infty$. Intuitively, one can guess that for large times the memory of the input state is washed out and that the leading contribution to the entropy comes from the Gaussian noise alone. Indeed it can be shown (see Methods) that, for every input state $\rho(0)$,

$$e^{S(\rho(t))/n} = \frac{et}{2} + \mathcal{O}(1) \,. \tag{19}$$

We then consider as input states the evolved $\rho_A(t_A)$ and $\rho_B(t_B)$, where we still have the freedom to let A and B evolve with different speeds by suitably choosing the dependence of their times $t_A(t)$ and $t_B(t)$ on a common time t, with the conditions:

$$t_A(0) = t_B(0) = 0 \,, \tag{20}$$

$$t_A, \ t_B \to \infty \quad \text{for} \quad t \to \infty \,. \tag{21}$$

From the composition laws of Gaussian channels, it follows that evolving ρ_A and ρ_B by times t_A and t_B before the application of the beam–splitter (or of the amplifier) produces at the output the state ρ_C evolved by a time

$$t_C = \lambda_A t_A + \lambda_B t_B \,. \tag{22}$$

The corresponding time dependent version of the qEPIs (5) and (7) can be rearranged in the following form:

$$1 \overset{?}{\geq} \frac{\lambda_A e^{S[\rho_A(t_A)]/n} + \lambda_B e^{S[\rho_B(t_B)]//n}}{e^{S[\rho_C(t_C)]/n}} . \tag{23}$$

Now if we plug in the asymptotic behavior (19), we see that the inequality is saturated for $t \to \infty$. The qEPI that we need to prove is simply (23) for $t = 0$ and this can be achieved if we are able to show that the RHS of (23) is monotonically increasing in time, *i.e.* that

$$\frac{d}{dt} \frac{\lambda_A e^{\frac{S_A}{n}} + \lambda_B e^{\frac{S_B}{n}}}{e^{\frac{S_C}{n}}} \overset{?}{\geq} 0 , \tag{24}$$

where we have put for simplicity

$$S_X = S[\rho_X(t_X)] \quad \text{for} \quad X = A, B, C . \tag{25}$$

From the *quantum de Bruijn identity* (16), the positivity of the derivative in (24) can be expressed as

$$\lambda_A e^{\frac{S_A}{n}} J_A \dot{t}_A + \lambda_B e^{\frac{S_B}{n}} J_B \dot{t}_B \overset{?}{\geq} \left(\lambda_A e^{\frac{S_A}{n}} + \lambda_B e^{\frac{S_B}{n}} \right) J_C \dot{t}_C . \tag{26}$$

Now we make use of the freedom that we have in choosing the functions $t_A(t)$, $t_B(t)$ and we impose them to satisfy the differential equation

$$\dot{t}_X = e^{S(t_X)/n} , \quad X = A, B , \tag{27}$$

with initial condition

$$t_X(0) = 0 . \tag{28}$$

Since the entropy is nonnegative, $\dot{t}_X \geq 1$ and (21) is satisfied. From Eq. (22) we have $\dot{t}_C = \lambda_A \dot{t}_A + \lambda_B \dot{t}_B$ and so the condition (26) reduces to

$$\left(\lambda_A e^{\frac{S_A}{n}} + \lambda_B e^{\frac{S_B}{n}} \right)^2 J_C \overset{?}{\leq} \lambda_A e^{\frac{2S_A}{n}} J_A + \lambda_B e^{\frac{2S_B}{n}} J_B . \tag{29}$$

At this point our quantum version of the *Stam inequality* (9) comes into play providing a useful upper bound to J_C,

$$J_C \leq \frac{J_A J_B}{\lambda_A J_B + \lambda_B J_A} . \tag{30}$$

By plugging it into (29) and rearranging the terms we get

$$\frac{\lambda_A \lambda_B \left(J_A e^{S_A/n} - J_B e^{S_B/n} \right)^2}{\lambda_A J_B + \lambda_B J_A} \geq 0 , \tag{31}$$

which is trivially satisfied because of the non-negativity of the Fisher information. This concludes the proof of both inequalities (5) and (7) and we can now focus on some of their direct implications.

C.   Linear inequalities

One of the features of the qEPI is that it is a significantly strong bound. For example from the concavity of the logarithm we directly get from (5) and (7) the respective linear relations, *i.e.*

$$S(\rho_C) \geq \lambda S(\rho_A) + (1 - \lambda)S(\rho_B), \tag{32}$$

$$S(\rho_C) \geq \frac{\kappa S(\rho_A) + (\kappa - 1)S(\rho_B)}{2\kappa - 1} + \ln(2\kappa - 1). \tag{33}$$

In the classical setting, the analogue of the first of these expressions is known to be formally equivalent to Eq. (2). For the quantum case however, such correspondence is no longer valid and Eqs. (32) and (33) appear to be weaker than (5) and (7), respectively. We remind also that Eq. (32) was originally conjectured in 10 and proven by König and Smith in 17 for all $\lambda \in [0, 1]$.

a.   Bound on the EPnI.–   It turns out that Eq. (5) is not the only way of generalizing the classical inequality (2). Another possible generalization was proposed and conjectured in Ref. 10 and 11. This is the *Entropy Photon number Inequality* (EPnI):

$$N(\rho_C) \overset{?}{\geq} \lambda N(\rho_A) + (1 - \lambda) N(\rho_B), \tag{34}$$

where $g(N) = (N + 1)\ln(N + 1) - N \ln N$ is the entropy of a single mode thermal state with mean photon number N, and $N(\rho) = g^{-1}(S(\rho)/n)$ is the mean photon number per mode of an n–mode thermal state with the same entropy of ρ. The EPnI states that fixing the input entropies S_A, S_B, the output entropy S_C is minimum when the inputs are thermal. Since the qEPI (5) is *not* saturated by thermal states (unless they have the same entropy), it is weaker than (and it is actually implied by) the EPnI (34), so our proof of qEPI does not imply the EPnI, which still remains an open conjecture. However, as we are going to show, the validity of the qEPI imposes a very tight bound (of the order of 0.132) on the maximum allowed violation of the EPnI (34).

The map $e^{S(\rho)/n} \mapsto N(\rho)$ from the entropy power to the entropy photon-number is the function $f(x) \equiv g^{-1}(\ln(x))$ defined on the interval $[1, \infty]$. Unfortunately it is convex and we cannot obtain the EPnI (34) from (5). Fortunately however, $f(x)$ is *not too convex* and is well approximated by a linear function. It is easy to show indeed that $f(x) = -1/2 + x/e + \delta(x)$, where $0 \leq \delta(x) \leq \delta(1) = 1/2 - 1/e \simeq 0.132$. This directly implies that the entropy photon

number inequality is valid up to such a small error,

$$N(\rho_C) - \lambda N(\rho_A) - (1 - \lambda)N(\rho_B) \geq 1/e - 1/2 \ . \tag{35}$$

As a side remark on the EPnI, we conjecture that an inequality similar to (34) should hold also in the case in which the mixing channel is the quantum amplifier,

$$N(\rho_C) \overset{?}{\geq} \kappa \ N(\rho_A) + (\kappa - 1) \ (N(\rho_B) + 1) \ , \tag{36}$$

but even in this case we do not have a proof.

D. Generalized minimum output entropy conjecture

Recently the so called *minimum output entropy conjecture* has been proved[14,15,25]. It claims (in the notation of this work) that when ρ_A is a Gaussian thermal state, the minimum output entropy $S(\rho_C)$ is achieved when the input ρ_B is the vacuum. The dual problem[10,11] is to fix $\rho_B = |0\rangle\langle 0|$ and to ask what is the minimum of $S(\rho_C)$ with the constraint that the input entropy is fixed $S(\rho_A) = \bar{S} > 0$. In Ref. 10 and 11 it was proved that the EPnI (34) implies that the minimum is achieved by the Gaussian centred thermal state with entropy $\bar{S}$, corresponding to an output entropy of $g\left(\lambda g^{-1}\left(\bar{S}\right)\right)$. Together with the EPnI, this *generalized conjecture* is still an open problem, however we can use our qEPI to obtain a tight lower bound on $S(\rho_C)$. The bound follows directly from (5) for $S(\rho_B) = 0$ and can be expressed as

$$S(\rho_C) \geq \ln\left[\lambda \ e^{\bar{S}} + (1 - \lambda)\right] \ . \tag{37}$$

The RHS of (37) is extremely close to the conjectured minimum $g\left(\lambda g^{-1}\left(\bar{S}\right)\right)$. Indeed the error between the two quantities $\Delta(\bar{S}, \lambda) = g\left(\lambda g^{-1}\left(\bar{S}\right)\right) - \ln\left[\lambda e^{\bar{S}} + (1 - \lambda)\right]$ is bounded by ~ 0.107 and moreover it decays to zero in large part of the parameter space $(\bar{S}, \lambda)$ (see Fig. 2). The plot in Fig. 2 provides also a useful hint about the small parameter region where a potential counter-example disproving the conjecture should be looked for.

Our qEPI (5), and in particular inequality (37), are also useful for bounding the capacity region of the bosonic broadcast channel. As explicitly discussed in the Methods, this bound is very close to the optimal one[12,13], which however relies on the still unproven conjecture[10,11] mentioned above.
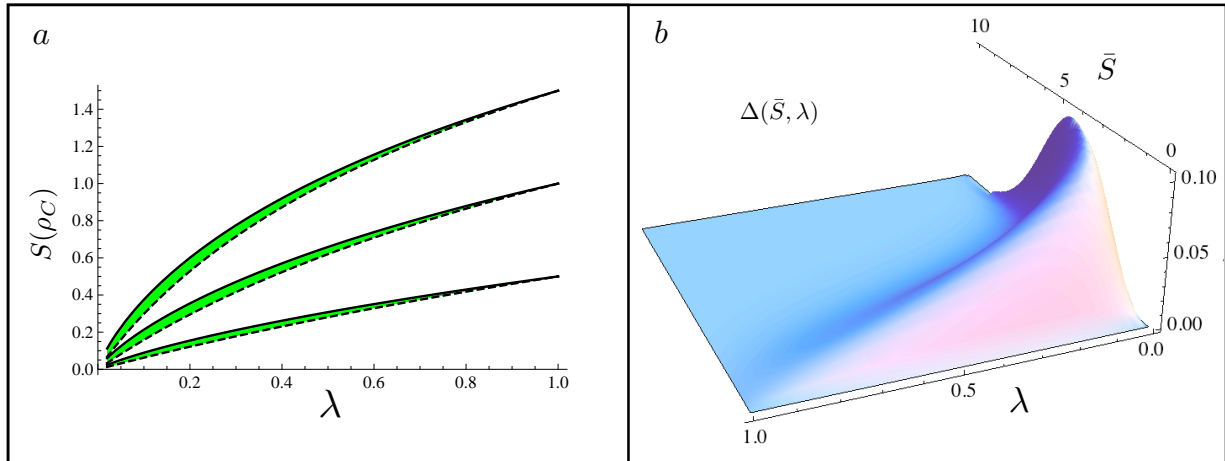
FIG. 2. **a** Plot of the output entropies as functions of λ and for different input entropies $\bar{S} = 0.5, 1, 1.5$. In full lines are the entropy achievable with a Gaussian input state while the dotted lines represent the lower bound (37). The corresponding minimum output entropies are necessarily constrained within the green regions. Notice that larger values of input entropies $\bar{S}$ are not considered in this plot because the Gaussian ansatz and the bound becomes practically indistinguishable. **b** Maximum allowed violation $\Delta(\bar{S}, \lambda)$ of the generalized minimum output entropy conjecture. The two axes are the input entropy $\bar{S}$ and the beam–splitter transmissivity λ. It is evident that the a potential violation of the conjecture is necessarily localized in the parameter space.

III. DISCUSSION

Understanding the complex physics of continuous variable quantum systems[20] represents a fundamental challenge of modern science which is crucial for developing an information technology capable of taking full advantage of quantum effects[21,26]. This task appears now to be within our grasp due to a series of very recent works which have solved a collection of long standing conjectures. Specifically, the minimum output entropy and output majorization conjectures (proposed in Ref. 16 and solved in Ref.s 14 and 25 respectively), the optimal Gaussian ensemble and the additivity conjecture (proposed in 27 and solved in Ref. 14), the optimality of Gaussian decomposition in the calculation of entanglement of formation[28] and of Guassian discord[29,30] for two-mode gaussian states (both solved in Ref. 15), the proof of the strong converse of the classical capacity theorem[31]. Our work represents a fundamental further step in this direction by extending the proof of 17 for the qEPI conjecture to include

11

all beam splitter transmissivities and by generalizing it to active bosonic transformations (e.g. amplification processes).

IV. METHODS

A. Proof of inequality (12)

In this section we will prove the inequality (12)

$$w_C^2 J_C \le w_A^2 J_A + w_B^2 J_B \tag{38}$$

for the quantum amplifier. Since the proof is analogue to the beam–splitter case, for clarity we present both.

We start recalling some basic characteristics of these channels. Their n output modes have annihilation operators

$$c_i = \sqrt{\lambda_A}\, a_i + \sqrt{\lambda_B}\, b_i \quad i = 1, \ldots, n \quad \text{(beam–splitter)} , \tag{39}$$

$$c_i = \sqrt{\lambda_A}\, a_i + \sqrt{\lambda_B}\, b_i^\dagger \quad i = 1, \ldots, n \quad \text{(amplifier)} , \tag{40}$$

with λ_A, $\lambda_B \ge 0$ such that

$$\lambda_A + \lambda_B = 1 \quad \text{(beam–splitter)} , \tag{41}$$

$$\lambda_A - \lambda_B = 1 \quad \text{(amplifier)} . \tag{42}$$

Let T be the time reversal matrix acting on the phase space, which reverses the signs of the quadratures P_i and satisfies

$$T = T^t = T^{-1} . \tag{43}$$

Let γ_A and γ_B be the covariance matrices of the two inputs; then the output will have a covariance matrix

$$\gamma_C = \lambda_A \gamma_A + \lambda_B \gamma_B \qquad \text{(beam–splitter)} , \tag{44}$$

$$\gamma_C = \lambda_A \gamma_A + \lambda_B T \gamma_B T \qquad \text{(amplifier)} . \tag{45}$$

For the displacement vectors we have instead

$$d_C = \sqrt{\lambda_A}\, d_A + \sqrt{\lambda_B}\, d_B \qquad \text{(beam–splitter)} , \tag{46}$$

$$d_C = \sqrt{\lambda_A}\, d_A + \sqrt{\lambda_B}\, T d_B \qquad \text{(amplifier)} . \tag{47}$$

1.   Compatibility with the Liouvillian

We recall **Lemma III.1** of 17: for any $t \geq 0$, the CPTPM $e^{t\mathcal{L}}$ is a Gaussian map acting on covariance matrices and displacement vectors by

$$
\begin{aligned}
\gamma \mapsto \gamma' &= \gamma + t\mathbb{1}_{2n} \, , \\
d \mapsto d' &= \quad d \, .
\end{aligned}
\tag{48}
$$

Then, if we choose as inputs the states $\rho_A(t_A)$ and $\rho_B(t_B)$ evolved with times t_A and t_B, the output will have covariance matrix and displacement vector

$$
\gamma_C(t) = \lambda_A \gamma_A + \lambda_B \gamma_B + \lambda_A t_A \mathbb{1}_{2n} + \lambda_B t_B \mathbb{1}_{2n} =
$$

$$
= \gamma_C(0) + t_C \mathbb{1}_{2n} \, ,
\tag{49}
$$

$$
d_C(t) = \sqrt{\lambda_A} \, d_A + \sqrt{\lambda_B} \, d_B = d_C(0) \, ,
\tag{50}
$$

in the case of the beam–splitter, and

$$
\gamma_C(t) = \lambda_A \gamma_A + \lambda_B T \gamma_B T + \lambda_A t_A \mathbb{1}_{2n} + \lambda_B t_B \mathbb{1}_{2n} =
$$

$$
= \gamma_C(0) + t_C \mathbb{1}_{2n} \, ,
\tag{51}
$$

$$
d_C(t) = \sqrt{\lambda_A} \, d_A + \sqrt{\lambda_B} \, T d_B = d_C(0) \, ,
\tag{52}
$$

in the case of the amplifier, where we have used $T^2 = \mathbb{1}_{2n}$ and we have put

$$
t_C = \lambda_A t_A + \lambda_B t_B \, .
\tag{53}
$$

Then, first evolving the inputs with times t_A, t_B and then applying the beam-splitter / amplifier is the same as first applying the beam–splitter/amplifier and then evolving the output with time t_C as in (53).

2.   Properties of quantum Fisher information

To prove inequality (38), we will follow the proof of the beam–splitter version for $\lambda = \frac{1}{2}$ in 17. First, given a smooth parametric family of states $\theta \mapsto \rho^{(\theta)}$, one can define (see Eq. (69) in 17) the associated quantum Fisher information with

$$
J\left(\rho^{(\theta)}; \theta\right) \equiv \frac{d^2}{d\theta^2} S\left(\rho^{(0)} \middle\| \rho^{(\theta)}\right)\bigg|_{\theta=0} \, .
\tag{54}
$$

It is linear in the parameter (17 **Lemma IV.1**):

$$J\left(\rho^{(c\theta)};\theta\right) = c^2 J\left(\rho^{(\theta)};\theta\right) , \tag{55}$$

and additive on product states (17 **Lemma IV.3**):

$$J\left(\rho_A^{(\theta)} \otimes \rho_B^{(\theta)};\theta\right) = J\left(\rho_A^{(\theta)};\theta\right) + J\left(\rho_B^{(\theta)};\theta\right) . \tag{56}$$

It is also always nonnegative (17 **Lemma IV.2**):

$$J\left(\rho^{(\theta)};\theta\right) \geq 0 , \tag{57}$$

and vanishes for $\theta = 0$, where it has a minimum. Then the data processing inequality for the relative entropy

$$S(\mathcal{E}(\hat{\rho})\|\mathcal{E}(\hat{\sigma})) \leq S(\hat{\rho}\|\hat{\sigma}) \tag{58}$$

implies that the quantum Fisher information is non–increasing under the application of any CPTP map $\mathcal{E}$ (17 **Theorem IV.4**):

$$J\left(\mathcal{E}\left(\rho^{(\theta)}\right);\theta\right) \leq J\left(\rho^{(\theta)};\theta\right) . \tag{59}$$

If the family is generated by conjugation with an exponential as in formula (76) of 17:

$$\rho^{(\theta)} = e^{i\theta H}\rho^{(0)}e^{-i\theta H} , \tag{60}$$

then (17 **Lemma IV.5**)

$$J\left(\rho^{(\theta)};\theta\right) = \text{Tr}\left(\rho^{(0)}\left[H,\left[H,\ln\rho^{(0)}\right]\right]\right) =$$
$$= \text{Tr}\left(\left[H,\left[H,\rho^{(0)}\right]\right]\ln\rho^{(0)}\right) . \tag{61}$$

For $R \in \{Q_j, P_j\}$ we define the displacement operator in the direction R as in 17, formula (79):

$$D_R(\theta) = \begin{cases} e^{i\theta P_j} & \text{if } R = Q_j , \\ e^{-i\theta Q_j} & \text{if } R = P_j . \end{cases} \tag{62}$$

For a state ρ, we consider the family of translated states

$$\rho^{(\theta, R)} = D_R(\theta)\rho D_R(\theta)^\dagger , \tag{63}$$

and its Fisher information $J\left(\rho^{(\theta,\,R)};\theta\right)$. We define the quantity $J(\rho)$ as the sum of the quantum Fisher information along all the phase space directions:

$$J(\rho) \equiv \sum_{k=1}^{2n} J\left(\rho^{(\theta,R_k)};\theta\right) \; . \tag{64}$$

Using (61), we get

$$J(\rho) = \sum_{i=1}^{n} \mathrm{Tr}\left(\left(\left[P_i,\left[P_i,\rho^{(0)}\right]\right] + \right.\right.$$
$$\left.\left. + \left[Q_i,\left[Q_i,\rho^{(0)}\right]\right]\right) \ln \rho^{(0)}\right) \; , \tag{65}$$

and since

$$\left.\frac{d}{dt}S\left(e^{t\mathcal{L}}\rho\right)\right|_{t=0} = -\mathrm{Tr}\left(\mathcal{L}(\rho)\ln\rho\right) \; , \tag{66}$$

we finally get

$$\frac{dS(\rho(t))}{dt} = \frac{1}{4}J(\rho) \; , \tag{67}$$

as in **Theorem V.1** of 17. The key point here is that if we define $\widetilde{J}$ with the time inverted quadratures

$$\widetilde{J}(\rho) \equiv \sum_{k=1}^{2n} J\left(\rho^{(\theta,TR_k)};\theta\right) \; , \tag{68}$$

the two definitions coincide:

$$\widetilde{J}(\rho) = J(\rho) \; , \tag{69}$$

since the P_j appear always quadratically in (65).

We now want to apply the data processing inequality (59) to our beam–splitter/amplifier channel to obtain the quantum Fisher information inequality.

3. Compatibility with translations

Let $\mathcal{E}$ be the channel associated with the beam–splitter/amplifier. Then

$$\mathcal{E}\left(\rho_A^{(w_A\theta,R)} \otimes \rho_B^{(w_B\theta,R)}\right) = \mathcal{E}\left(\rho_A \otimes \rho_B\right)^{(w_C\theta,R)} \; , \tag{70}$$

for the beam–splitter, and

$$\mathcal{E}\left(\rho_A^{(w_A\theta,R)} \otimes \rho_B^{(w_B\theta,TR)}\right) = \mathcal{E}\left(\rho_A \otimes \rho_B\right)^{(w_C\theta,R)} \; , \tag{71}$$

15

for the amplifier, *i.e.* translating the inputs by

$$d_A = w_A \theta d_R \quad d_B = w_B \theta d_R \qquad \text{(beam–splitter)} , \qquad (72)$$

$$d_A = w_A \theta d_R \quad d_B = w_B \theta T d_R \qquad \text{(amplifier)} , \qquad (73)$$

(notice the time reversal) and then applying the beam–splitter/amplifier is the same as applying the beam–splitter/amplifier and translating the output by $w_C \theta d_R$, where d_R is the phase space vector associated to the operator R and

$$w_C = \sqrt{\lambda_A} \, w_A + \sqrt{\lambda_B} \, w_B . \qquad (74)$$

The proof follows straightforwardly evaluating the displacement vectors with (46), (47): if we translate the inputs and then apply the beam–splitter we have

$$d_A(\theta) = d_A + w_A \theta d_R , \qquad (75)$$

$$d_B(\theta) = d_B + w_B \theta d_R , \qquad (76)$$

$$\begin{aligned} d_C(\theta) &= \sqrt{\lambda_A} \, d_A(\theta) + \sqrt{\lambda_B} \, d_B(\theta) = \\ &= \sqrt{\lambda_A} \, d_A + \sqrt{\lambda_B} \, d_B + \sqrt{\lambda_A} \, w_A \theta d_R + \\ &\quad + \sqrt{\lambda_B} \, w_B \theta d_R = \\ &= d_C(0) + w_C \theta d_R , \end{aligned} \qquad (77)$$

which is what we would get translating the output by $w_C \theta d_R$. The same happens for the amplifier:

$$d_A(\theta) = d_A + w_A \theta d_R , \qquad (78)$$

$$d_B(\theta) = d_B + w_B \theta T d_R , \qquad (79)$$

$$\begin{aligned} d_C(\theta) &= \sqrt{\lambda_A} \, d_A(\theta) + \sqrt{\lambda_B} \, T d_B(\theta) = \\ &= \sqrt{\lambda_A} \, d_A + \sqrt{\lambda_B} \, T d_B + \sqrt{\lambda_A} \, w_A \theta d_R + \\ &\quad + \sqrt{\lambda_B} \, w_B \theta d_R = \\ &= d_C(0) + w_C \theta d_R . \end{aligned} \qquad (80)$$

Now we can apply the data processing inequality (59) to (70) and (71). Using the additivity (56) and the linearity (55) of the Fisher information, we get

$$w_C^2 J \left(\rho_C^{(\theta,R)}; \theta \right) \leq w_A^2 J \left(\rho_A^{(\theta,R)}; \theta \right) + w_B^2 J \left(\rho^{(\theta,R)}; \theta \right) , \qquad (81)$$

for the beam–splitter, and

$$w_C^2 J\left(\rho_C^{(\theta,R)};\theta\right) \le w_A^2 J\left(\rho_A^{(\theta,R)};\theta\right) + w_B^2 J\left(\rho^{(\theta,TR)};\theta\right) , \tag{82}$$

for the amplifier. These two results are identical, apart from the time reversal in B in the amplifier case. Finally, summing over the phase space direction we get in both cases the desired inequality

$$w_C^2 J_C \le w_A^2 J_A + w_B^2 J_B , \tag{83}$$

since we have proved in (69) that the time reversal does not affect the sum.

B. Proof of the asymptotic scaling (19)

In 17, **Corollary III-4** it is shown that

$$\exp\left(\frac{1}{n}S\left(e^{t\mathcal{L}}\hat\rho\right)\right) \ge \frac{et}{2} + \mathcal{O}(1) ; \tag{84}$$

here we prove the upper bound.

Let $\hat\rho_G$ be the Gaussianized version of $\hat\rho$, *i.e.* the Gaussian state with the same first and second moments. Since Gaussianization always increases entropy[32] and commutes with the Liouvillean $\mathcal{L}$[17], $S\left(e^{t\mathcal{L}}\hat\rho\right) \le S\left(e^{t\mathcal{L}}\hat\rho_G\right)$. The covariance matrix of $e^{t\mathcal{L}}\hat\rho$ and $e^{t\mathcal{L}}\hat\rho_G$ is (17, **Lemma III.1**) $\sigma + t\mathbb{1}_{2n}$, where σ is the one of $\hat\rho$. Let λ_0 be the maximum eigenvalue of σ. Then $\sigma + t\mathbb{1}_{2n} \le (\lambda_0 + t)\mathbb{1}_{2n}$, i.e. the Gaussian thermal state with covariance matrix $(\lambda_0 + t)\mathbb{1}_{2n}$ can be obtained adding (non–white) Gaussian noise to $e^{t\mathcal{L}}\hat\rho_G$. Since the additive noise channel is unital, it always increases the entropy, and

$$S\left(e^{t\mathcal{L}}\hat\rho_G\right) \le ng\left(\frac{\lambda_0 + t - 1}{2}\right) . \tag{85}$$

Since

$$g\left(x - \frac{1}{2}\right) = \ln(ex) + \mathcal{O}\left(\frac{1}{x^2}\right) \text{ for } x \to \infty , \tag{86}$$

putting all together we get

$$\exp\left(\frac{1}{n}S\left(e^{t\mathcal{L}}\hat\rho\right)\right) \le \frac{et}{2} + \mathcal{O}(1) . \tag{87}$$

17

C. Capacity region of the bosonic broadcast channel

In Ref. 12 and 13 it is proven that, trusting the minimum output entropy conjecture of 10 and 11 (which is a particular case of the still unproven EPnI), the capacity region for a lossless bosonic broadcast channel is parametrically described by the inequalities

$$R_B \leq g\left(\lambda\beta\bar{N}\right) , \tag{88}$$
$$R_C \leq g\left((1-\lambda)\bar{N}\right) - g\left((1-\lambda)\beta\bar{N}\right) ,$$

with R_B and R_C representing the achievable communication rates the sender of the information can establish when signaling simultaneously to two independent receivers B and C, respectively, when coding his messages into a single bosonic mode which splits at a beam splitter of transmissivity $\lambda \geq \frac{1}{2}$ (the transmitted signals being routed to B and the reflected ones to C, see 12 and 13 for details). In this expression $\beta \in [0, 1]$ represents the fraction of the sender's average photon number that is meant to convey information to B, with the remainder to be used to communicate information to C. $\bar{N} \geq 0$ instead is the maximum average mean input photon number employed in the communication per channel uses. Our qEPI inequality (5) provides instead the weaker bound

$$R_B \leq g\left(\lambda\beta\bar{N}\right) , \tag{89}$$
$$R_C \leq g\left((1-\lambda)\bar{N}\right) - \ln\frac{(1-\lambda)e^{g\left(\lambda\beta\bar{N}\right)} + 2\lambda - 1}{\lambda} .$$

A comparison between Eq. (89) and the conjectured region (88) is shown in Fig. 3: the discrepancy being small.

REFERENCES

[1]Shannon C. E., A Mathematical Theory of Communication, Bell Syst. Tech. J. 27, 379-423 (1948).

[2]Dembo, A., Cover, T. & Thomas, J., Information theoretic inequalities. IEEE Trans. Inf. Theory 37, 1501-1518 (1991).

[3]Stam, A., Some inequalities satisfied by the quantities of information of Fisher and Shannon. Inform. Control 2, 101-112 (1959).

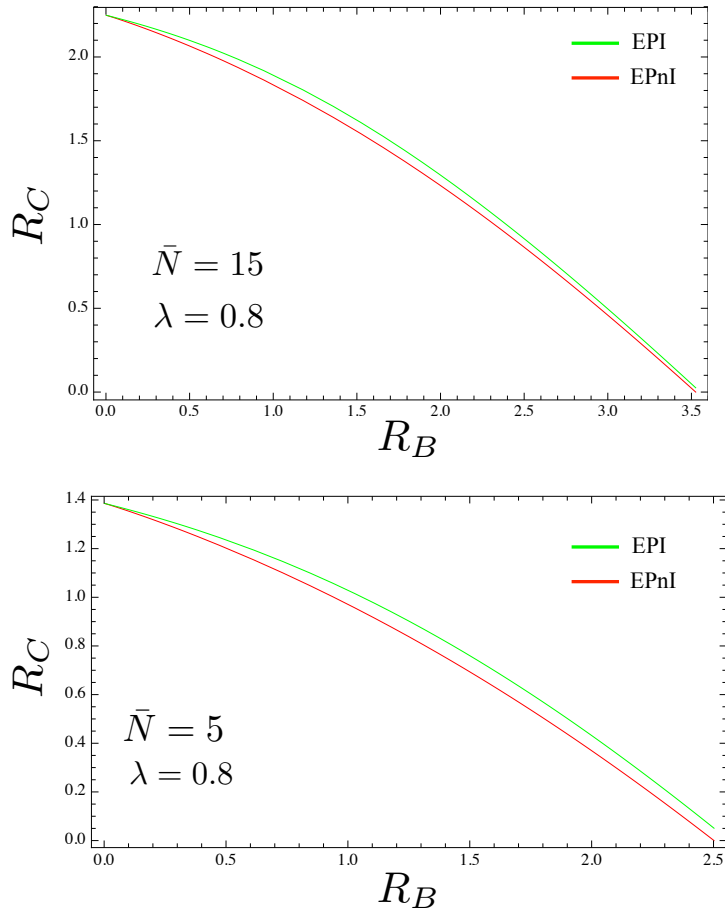[4]Blachman, N., The convolution inequality for entropy powers. IEEE Trans. Inf. Theory 11, 267-271 (1965).

FIG. 3. Capacity region (expressed in nats per channel uses) for a broadcasting channel[12,13] in which the sender is communicating simultaneously with two receivers (B and C) via a single bosonic mode which splits at a beam splitter of transmissivity λ (B receiving the transmitted signals, while C receiving the reflected one), under input energy constraint which limits the mean photon number of the input messages to be smaller than $\bar{N}$. The region delimited by the red curve represents the achievable rates R_B and R_C which would apply if the (still unproven) EPnI conjecture (34) held. The green curve instead is the bound one can derive via Eq. (37) from the EPI inequality (5) we have proven in this paper.

[5]Verdú, S. & Guo, D., A simple proof of the entropy-power inequality. IEEE Trans. Inf. Theory 52, 2165-2166 (2006).

[6]Rioul, O., Information theoretic proofs of entropy power inequalities. IEEE Trans. Inf. Theory 57, 33-55 (2011).

[7]Guo, D., Shamai, S. & Verdú, S. Proof of Entropy Power Inequalities Via MMSE in Proceedings of Information Theory 2006, IEEE International Symposium on Information Theory, 1011-1015 (2006).

[8]Bergmans, P., A simple converse for broadcast channels with additive white Gaussian noise (corresp.). IEEE Trans. Inf. Theory 20, 279-280 (1974).

[9]Leung-Yan-Cheong, S. & Hellman, M., The Gaussian wire-tap channel. IEEE Trans. Inf. Theory 24, 451-456 (1978).

[10]Guha, S., Erkmen, B.I. & Shapiro, J.H., The Entropy Photon-Number Inequality and its consequences. Information Theory and Applications Workshop, 128-130 (2008)

[11]Guha, S., Shapiro, J. & Erkmen, B., Capacity of the bosonic wiretap channel and the entropy photon-number inequality in Proceedings of Information Theory 2008, IEEE International Symposium on Information Theory, 91-95 (2008).

[12]Guha, S. & Shapiro, J., Classical Information Capacity of the Bosonic Broadcast Channel in Proceedings of Information Theory 2007, IEEE International Symposium on Information Theory, 1896-1900 (2007).

[13]Guha S., Shapiro, J. H. & Erkmen, B. I., Classical capacity of bosonic broadcast communication and a minimum output entropy conjecture. Phys. Rev. A 76, 032303 (2007).

[14]Giovannetti V., Holevo, A. S. & García-Patrón, R., A solution of the Gaussian optimizer conjecture. Preprint at http://lanl.arxiv.org/abs/1312.2251 (2013).

[15]Giovannetti V., García-Patrón R., Cerf, N. J. & Holevo A. S., Ultimate communication capacity of quantum optical channels by solving the Gaussian minimum-entropy conjecture. Preprint at http://lanl.arxiv.org/abs/1312.6225 (2013).

[16]Giovannetti V., Guha, S., Lloyd, S., Maccone, L. & Shapiro, J. H., Minimum output entropy of bosonic channels: A conjecture. Phys. Rev. A 70, 032315 (2004).

[17]König R. & Smith G., The Entropy Power Inequality for Quantum Systems. IEEE Trans. Inf. Theory 60, 1536-1548 (2014).

[18]König R. and Smith G., Limits on classical communication from quantum entropy power inequalities. Nature Photon. 7, 142-146 (2013).

[19]Barron, A. R., Entropy and the Central Limit Theorem. Ann. Prob. 14, 336-342 (1986).

[20]Braunstein S. L. and van Loock P., Quantum information with continuous variables. Rev. Mod. Phys. 77, 513-577 (2005).

[21]Weedbroock C. et al., Gaussian Quantum Information. Rev. Mod. Phys. 84, 621-669 (2012).

[22]Walls D. F. & Milburn G. J., Quantum Optics (Springer, 1994).

[23]Holevo, A. S., Quantum systems, channels, information: a mathematical introduction, (Walter de Gruyter, 2012).

[24]Kagan, A. & Yu, T., Some inequalities related to the Stam inequality. Appl. Math. 53, 195-205 (2008).

[25]Mari, A., Giovannetti V. & Holevo, A. S., Quantum state majorization at the output of bosonic Gaussian channels. Nature Commun. 5, 3826 (2014).

[26]Caves C. M. & Drummond P. B., Quantum limits on bosonic communication rates, Rev. Mod. Phys. 66, 481-537 (1994).

[27]Holevo A. S. & Werner R. F., Evaluating capacities of bosonic Gaus- sian channels, Phys. Rev. A 63, 032312 (2001).

[28]Giedke G., Wolf, M. M., Krüger, O., Werner, R. F. & Cirac, J. I., Entanglement of Formation for Symmetric Gaussian States, Phys. Rev. Lett. 91, 107901 (2003).

[29]Modi, K., Brodutch, A., Cable, H., Paterek, T. & Vedral, V., The classical-quantum boundary for correlations: Discord and related measures. Rev. Mod. Phys. 84, 1655-1707 (2012).

[30]Pirandola, S., Cerf, N. J., Braunstein, S. L. & Lloyd, S., Optimality of Gaussian discord. Preprint at http://lanl.arxiv.org/abs/1309.2215 (2013).

[31]Bardhan, B. R., García-Patrón, R., Wilde, M. M. & Winter, A., Strong converse for the classical capacity of optical quantum communication channels. Preprint at http://lanl.arxiv.org/abs/1401.4161 (2014).

[32]M. M. Wolf, G. Giedke, & J. I. Cirac, Extremality of Gaussian quantum states, Phys. Rev. Lett. 96, 080502 (2006).

V. ACKNOWLEDGEMENTS

VI. SUPPLEMENTARY MATERIAL

A. Linear inequality for the quantum amplifier

The proof of the inequality (33)

$$S(\rho_C) \geq \frac{\kappa S(\rho_A) + (\kappa - 1)S(\rho_B)}{2\kappa - 1} + \ln(2\kappa - 1) \ , \tag{90}$$

for the amplifier is straightforward: the entropy power inequality (7) can be rewritten as

$$S_C \geq \ln\left(\frac{\lambda_A}{\lambda_A + \lambda_B}e^{S_A} + \frac{\lambda_B}{\lambda_A + \lambda_B}e^{S_B}\right) + \ln(\lambda_A + \lambda_B) \ , \tag{91}$$

which for the concavity of the logarithm implies

$$S_C \geq \frac{\lambda_A S_A + \lambda_B S_B}{\lambda_A + \lambda_B} + \ln(\lambda_A + \lambda_B) \ , \tag{92}$$

i.e. (33).

This result can also be proven without recurring to the qEPI: let us evolve with the Liouvillian the inputs with equal times

$$t_A = t_B = t \ . \tag{93}$$

The corresponding evolution time for the output will be

$$t_C = (\lambda_A + \lambda_B)t \ . \tag{94}$$

Recalling the asymptotic behaviour of the entropies (19), both sides of (33) behave as

$$\ln\frac{et}{2} + \ln(\lambda_A + \lambda_B) + \mathcal{O}\left(\frac{1}{t}\right) \tag{95}$$

for $t \to \infty$, and (33) is asymptotically saturated. Then we have only to check that

$$\frac{d}{dt}S_C \leq \frac{d}{dt}\frac{\lambda_A S_A + \lambda_B S_B}{\lambda_A + \lambda_B} \ , \tag{96}$$

i.e. that

$$J_C \leq \frac{\lambda_A J_A + \lambda_B J_B}{(\lambda_A + \lambda_B)^2} \ , \tag{97}$$

where we have used

$$\frac{d}{dt}S_C = (\lambda_A + \lambda_B)J_C \ . \tag{98}$$

But (97) is exactly what we get if we plug in the quantum Fisher information inequality (38)

$$w_A = \sqrt{\lambda_A} \ , \tag{99}$$

$$w_B = \sqrt{\lambda_B} \ . \tag{100}$$

22

B. Bound on EPnI

We want to evaluate how close is qEPI (5) to EPnI (34) and prove (35). (5) implies for the output entropy photon number

$$N_C \geq g^{-1}\left(\ln\left(\lambda_A e^{g(N_A)} + \lambda_B e^{g(N_B)}\right)\right) . \tag{101}$$

(34) is stronger than (5), and in fact

$$g^{-1}\left(\ln\left(\lambda_A e^{g(N_A)} + \lambda_B e^{g(N_B)}\right)\right) \leq \lambda_A N_A + \lambda_B N_B , \tag{102}$$

since the function $g^{-1}\left(\ln\left(x\right)\right)$ is increasing and convex. Since $e^{g(N)}$ for $N \to \infty$ goes like

$$e^{g(N)} = e\left(N + \frac{1}{2}\right) + \mathcal{O}\left(\frac{1}{N}\right) , \tag{103}$$

we have for $x \to \infty$

$$g^{-1}\left(\ln x\right) = \frac{x}{e} - \frac{1}{2} + \mathcal{O}\left(\frac{1}{x}\right) . \tag{104}$$

If we define as in the main text

$$\delta(x) \equiv g^{-1}\left(\ln x\right) - \frac{x}{e} + \frac{1}{2} , \tag{105}$$

δ is convex, decreasing and

$$\lim_{x \to \infty} \delta(x) = 0 . \tag{106}$$

We can also evaluate

$$\delta(1) = \frac{1}{2} - \frac{1}{e} , \tag{107}$$

and for any x_A, $x_B \geq 1$ we have

$$\delta(\lambda_A x_A + \lambda_B x_B) \geq \lambda_A \delta(x_A) + \lambda_B \delta(x_B) - \left(\frac{1}{2} - \frac{1}{e}\right) . \tag{108}$$

Since

$$g^{-1}\left(\ln\left(\lambda_A x_A + \lambda_B x_B\right)\right) -$$
$$- \lambda_A g^{-1}\left(\ln x_A\right) - \lambda_B g^{-1}\left(\ln x_B\right) =$$
$$= \delta\left(\lambda_A x_A + \lambda_B x_B\right) - \lambda_A \delta(x_A) - \lambda_B \delta(x_B) , \tag{109}$$

in the case $x_A = e^{S_A}$, $x_B = e^{S_B}$ we get

$$g^{-1}\left(\ln\left(\lambda_A e^{S_A} + \lambda_B e^{S_B}\right)\right) - \lambda_A N_A - \lambda_B N_B =$$
$$= \delta\left(\lambda_A e^{S_A} + \lambda_B e^{S_B}\right) - \lambda_A \delta(e^{S_A}) - \lambda_B \delta(e^{S_B}) , \tag{110}$$

and we can conclude from (101) that

$$N_C \geq \lambda_A N_A + \lambda_B N_B +$$

$$+ \delta(\lambda_A e^{S_A} + \lambda_B e^{S_B}) - \lambda_A \delta(e^{S_A}) - \lambda_B \delta(e^{S_B}) \geq$$

$$\geq \lambda_A N_A + \lambda_B N_B - \left(\frac{1}{2} - \frac{1}{e}\right) , \tag{111}$$

so the (34) violation can be at most

$$\frac{1}{2} - \frac{1}{e} \simeq 0.132 . \tag{112}$$