



PAPER

OPEN ACCESS

RECEIVED

8 November 2019

REVISED

19 February 2020

ACCEPTED FOR PUBLICATION

2 March 2020

PUBLISHED

16 April 2020

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Low-cost limit of classical communication with restricted quantum measurements

Ludwig Kunz^{1,2} , Marcin Jarzyna¹ , Wojciech Zwoleński^{1,2} and Konrad Banaszek^{1,2} ¹ Centre for Quantum Optical Technologies, Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097 Warszawa, Poland² Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, PolandE-mail: l.kunz@cent.uw.edu.pl**Keywords:** superadditivity, classical communication, structured receivers, low-power limit, photon-starved communication

Abstract

We consider a communication scenario where classical information is encoded in an ensemble of quantum states that admit a power series expansion in a cost parameter and converge to a single zero cost state with vanishing cost. For a given measurement scheme, we derive an approximate expression for mutual information in the leading order of the cost parameter. The general results are applied to selected problems in optical communication, where coherent states of light are used as input symbols and the cost is quantified as the average number of photons per symbol. We show that for an arbitrary individual measurement on phase shift keyed (PSK) symbols, the photon information efficiency is upper bounded by 2 nats of information per photon in the low-cost limit, which coincides with the conventional homodyne detection bound. The presented low-cost approximation facilitates a systematic analysis of few-symbol measurements that exhibit superadditivity of accessible information. For the binary PSK alphabet of coherent states, we present designs for two- and three-symbol measurement schemes based on linear optics, homodyning, and single photon detection that offer respectively 2.49% and 3.40% enhancement relative to individual measurements. We also show how designs for scalable superadditive measurement schemes emerge from the introduced low-cost formalism.

1. Introduction

The celebrated Holevo quantity [1] provides a powerful tool to identify how much classical information can be transmitted using symbols drawn from a given ensemble of quantum states. It specifies an upper bound on the mutual information that can be achieved with any quantum measurement satisfying physical constraints. In general, attaining the Holevo quantity requires a collective measurement on an arbitrary number of symbols [2–5]. When the available class of measurements is restricted, for example to individual or few-symbol detection, optimization of the mutual information with respect to the measurement strategy often becomes a rather challenging task for which universal methods are missing. Interestingly, the maximum mutual information that can be obtained using measurements on a restricted number of symbols, known as accessible information, exhibits superadditive behavior with the number of symbols [6, 7]. This is intimately related to the fact that quantum measurement reveals in general only partial knowledge about the measured states and more relevant information can be retrieved through collective detection strategies. Theoretical analysis of superadditivity of accessible information is highly nontrivial even for binary ensembles of elementary input states [8–10]. Another problem is the actual physical implementation of optimal measurements, e.g. in the optical domain.

The purpose of this paper is to present a systematic expansion of the mutual information in a cost parameter which characterizes the ensemble of quantum states used for communication. It is assumed that the quantum states admit a power series expansion in the cost parameter and converge to the same zero-cost state when the value of the cost parameter tends to zero. This allows us to derive an asymptotic bound on the mutual

information in the leading order of the cost parameter for a given measurement scheme. An essential advantage of the formalism developed here is that the characteristics of the ensemble enters the asymptotic expression through a set of state vectors that are independent of the cost parameter. As we demonstrate here, this greatly simplifies identification of bounds on mutual information attainable with few-symbol measurements and provides a systematic method to construct experimental schemes that demonstrate superadditivity of accessible information.

The usefulness of the presented approach is illustrated with a number of examples motivated by optical communication. In this context, a natural measure for the communication cost is the mean photon number per symbol. Our expansion provides non-trivial bounds on the photon information efficiency (PIE), which specifies the amount of information that can be transmitted in one photon. In particular, we show that when information is encoded in the phase of coherent states, the maximum attainable PIE is 2 nats per photon in the low-cost limit when arbitrary individual measurements are permitted. This coincides with the commonly used in optical communications reference value derived from the Shannon–Hartley theorem, which assumes the specific case of homodyne detection [11]. We also provide a relation between the attainable PIE and the peak-to-average power ratio of a general coherent state ensemble used for communication. Further, we investigate superadditivity of accessible information in the case of few-symbol measurements performed on the binary ensemble of coherent states with equal mean photon number and opposite phases. We show that linear optical setups combining photon counting and homodyne detection can be used to demonstrate superadditivity for coherent states carrying on average less than approx. 0.01 photon. We also discuss how scalable communication schemes exhibiting the superadditivity effect [10, 12] emerge from the presented formalism.

This paper is organized as follows. Section 2 presents the mathematical formulation of the problem. The asymptotic expansion of mutual information in the limit of the vanishing cost is carried out in section 3. The results are applied to communication with an ensemble of coherent states and individual measurements in section 4. The case of few-symbol measurements on the binary ensemble of coherent states is discussed in section 5. Finally, section 6 concludes the paper.

2. Problem formulation

A communication scheme can be regarded as encoding messages onto quantum states of certain physical carriers which after transmission are subsequently measured at the receiver. The measurement results are processed to decode the input message. We shall consider classical information encoded in a finite ensemble of quantum states, labeled using an index j and described, after transmission, by density operators $\hat{\rho}_j(\zeta)$, with respective input probabilities p_j . The states depend on a real nonnegative parameter ζ characterizing their cost. Physically, the cost may correspond e.g. to the mean photon number per symbol in the case of optical communication. We will assume that each $\hat{\rho}_j(\zeta)$ is an analytical operator-valued function of ζ and that in the limit $\zeta \rightarrow 0$ all states $\hat{\rho}_j(\zeta)$ converge to the same state $\hat{\rho}^{(0)}$ which we shall call the *zero-cost state*. Furthermore, we will assume that the states $\hat{\rho}_j(\zeta)$ admit a power series expansion in the cost parameter around $\zeta = 0$. The asymptotic analysis will be based on an expansion up to the second order:

$$\hat{\rho}_j(\zeta) \approx \hat{\rho}^{(0)} + \zeta \hat{\rho}_j^{(1)} + \zeta^2 \hat{\rho}_j^{(2)}, \quad (1)$$

where

$$\hat{\rho}_j^{(1)} = \left. \frac{\partial \hat{\rho}_j}{\partial \zeta} \right|_{\zeta=0}, \quad \hat{\rho}_j^{(2)} = \left. \frac{1}{2} \frac{\partial^2 \hat{\rho}_j}{\partial \zeta^2} \right|_{\zeta=0}. \quad (2)$$

Note that the zeroth-order term is independent of j as it is given by the zero-cost state. The above expansion is assumed to be valid for sufficiently small values of the cost parameter ζ .

The most general measurement that can be performed on a quantum system is described by a positive operator-valued measure (POVM) with elements \hat{Q}_r that are positive semidefinite and add up to the identity operator, $\sum_r \hat{Q}_r = \hat{1}$. The index r specifies the measurement result. For simplicity we consider here a finite set of possible measurement outcomes. The conditional probability $p_{r|j}(\zeta)$ of obtaining an outcome r given the j th input state is defined by Born's rule

$$p_{r|j}(\zeta) = \text{Tr}[\hat{Q}_r \hat{\rho}_j(\zeta)]. \quad (3)$$

The marginal probability of the r th outcome for the ensemble is $p_r(\zeta) = \sum_j p_j p_{r|j}(\zeta)$.

The maximum amount of classical information that can be transmitted using the input ensemble of quantum states $\hat{\rho}_j$ with respective probabilities p_j and a measurement \hat{Q}_r is given by mutual information I calculated for the joint probability distribution $p_j p_{r|j}(\zeta)$, where the conditional probabilities $p_{r|j}(\zeta)$ are given by equation (3). It will be convenient to decompose mutual information into a sum of contributions from

individual measurement results,

$$I = \sum_r I_r \quad (4)$$

with I_r given by

$$I_r = \sum_j p_j p_{r|j}(\zeta) \log \left(\frac{p_{r|j}(\zeta)}{p_r(\zeta)} \right). \quad (5)$$

Throughout this paper we use natural logarithms and the information is measured in nats, $1 \text{ nat} = \log_2 e$ bits.

Our overall objective will be to optimize mutual information over measurement operators chosen from a restricted class and input probabilities for a given set (constellation) of symbols $\varrho_j(\zeta)$ in the regime of a small cost parameter. The first step, presented in the next section, will be a systematic expansion of the mutual information I in the cost parameter ζ that will provide an upper bound on the leading-order term. From now on, it is assumed that the measurement operators \hat{Q}_r do not depend on the cost parameter. This restricted scenario is motivated by practical optical communication that uses direct detection or homodyne detection. We will see that even under this restriction the low-cost analysis of mutual information yields non-trivial results regarding the superadditivity of accessible information. In section 5.1 we will present an example of further enhancement when the measurement can be adjusted to the value of the cost parameter.

3. Asymptotic expansion

In the limit of the vanishing cost parameter, equation (5) can be expressed through expansion of the conditional probabilities $p_{r|j}(\zeta)$ in ζ , which also determines marginal probabilities $p_r(\zeta)$. As for $\zeta \rightarrow 0$ all the states $\hat{\varrho}_j(\zeta)$ converge to the zero-cost state $\hat{\varrho}^{(0)}$, it will be convenient to introduce a shorthand notation

$$p_r^{(0)} = \text{Tr}[\hat{Q}_r \hat{\varrho}^{(0)}] = p_{r|j}(0) = p_r(0). \quad (6)$$

The expansion given in equation (1) implies the following approximate expressions for conditional probabilities

$$p_{r|j}(\zeta) \approx p_r^{(0)} + \zeta p_{r|j}^{(1)} + \zeta^2 p_{r|j}^{(2)}, \quad (7)$$

where $p_{r|j}^{(k)} = \text{Tr}[\hat{Q}_r \hat{\varrho}_j^{(k)}]$, $k = 1, 2$. The marginal probabilities are correspondingly approximated by

$$p_r(\zeta) \approx p_r^{(0)} + \zeta p_r^{(1)} + \zeta^2 p_r^{(2)} \quad (8)$$

with $p_r^{(k)} = \sum_j p_j p_{r|j}^{(k)}$, where $k = 1, 2$. Equivalently, one can write $p_r^{(k)} = \text{Tr}[\hat{Q}_r \hat{\varrho}_{\text{ens}}^{(k)}]$, where the operators $\hat{\varrho}_{\text{ens}}^{(k)}$ are given by convex combinations

$$\hat{\varrho}_{\text{ens}}^{(k)} = \sum_j p_j \hat{\varrho}_j^{(k)}, \quad k = 1, 2 \quad (9)$$

of derivatives defined in equation (2). It should be noted that the zeroth-order term for both conditional probabilities in equation (7) and marginal probabilities in equation (8) is identical.

Further steps for a given measurement operator \hat{Q}_r depend on whether the outcome r can be generated by the zero-cost state or not. In the former case, corresponding to the mathematical condition $p_r^{(0)} = \text{Tr}[\hat{Q}_r \hat{\varrho}^{(0)}] > 0$, we will call \hat{Q}_r a *type- \mathcal{Z}* operator and use \mathcal{Z} to denote the corresponding set of measurement outcomes. Measurement operators for which $p_r^{(0)} = \text{Tr}[\hat{Q}_r \hat{\varrho}^{(0)}] = 0$ will be referred to as *type- \mathcal{Z}^\perp* operators and the set of measurement outcomes produced by these operators will be denoted by \mathcal{Z}^\perp . We will decompose

$$I = \sum_{r \in \mathcal{Z}} I_r + \sum_{r \in \mathcal{Z}^\perp} I_r \quad (10)$$

and analyze separately contributions to mutual information from the two sets of measurement outcomes.

In the case of type- \mathcal{Z} operators, the logarithm in equation (5) can be expanded into a power series in the cost parameter ζ , which yields in the leading order

$$I_r \approx \tilde{I}_r = \frac{\zeta^2}{2p_r^{(0)}} \sum_j p_j (p_{r|j}^{(1)} - p_r^{(1)})^2. \quad (11)$$

The difference $p_{r|j}^{(1)} - p_r^{(1)}$ appearing under summation can be written as

$$p_{r|j}^{(1)} - p_r^{(1)} = \text{Tr}[\hat{Q}_r (\hat{\varrho}_j^{(1)} - \hat{\varrho}_{\text{ens}}^{(1)})]. \quad (12)$$

We will transform this expression further using techniques borrowed from quantum estimation theory [13–16]. Let us recall the notion of the symmetric logarithmic derivative (SLD), defined implicitly at $\zeta = 0$ for individual

and ensemble density operators by respective equations

$$\hat{\rho}_j^{(1)} = \left. \frac{\partial \hat{\rho}_j}{\partial \zeta} \right|_{\zeta=0} = \frac{1}{2}(\hat{L}_j \hat{\rho}^{(0)} + \hat{\rho}^{(0)} \hat{L}_j), \quad (13)$$

$$\hat{\rho}_{\text{ens}}^{(1)} = \left. \frac{\partial \hat{\rho}_{\text{ens}}}{\partial \zeta} \right|_{\zeta=0} = \frac{1}{2}(\hat{L}_{\text{ens}} \hat{\rho}^{(0)} + \hat{\rho}^{(0)} \hat{L}_{\text{ens}}), \quad (14)$$

that are assumed to be well defined [17, 18] for physical scenarios considered here. Subtracting the expression in equation (14) from equation (13) and inserting the difference into equation (12) yields

$$p_{r|j}^{(1)} - p_r^{(1)} = \frac{1}{2} \text{Tr}[\hat{Q}_r(\hat{D}_j \hat{\rho}^{(0)} + \hat{\rho}^{(0)} \hat{D}_j)], \quad (15)$$

where

$$\hat{D}_j = \hat{L}_j - \hat{L}_{\text{ens}}. \quad (16)$$

Using the Schwarz inequality

$$|\text{Tr}[\hat{A}\hat{B}]|^2 \leq \text{Tr}[\hat{A}^\dagger \hat{A}] \text{Tr}[\hat{B}^\dagger \hat{B}] \quad (17)$$

applied to the operators

$$\hat{A} = \frac{1}{\sqrt{p_r^{(0)}}} \sqrt{\hat{\rho}^{(0)}} \sqrt{\hat{Q}_r}, \quad \hat{B} = \sqrt{\hat{Q}_r} \hat{D}_j \sqrt{\hat{\rho}^{(0)}} \quad (18)$$

and their hermitian conjugates \hat{A}^\dagger and \hat{B}^\dagger and inserting the result into equation (11) yields the following upper bound on the mutual information contribution from the measurement outcome $r \in \mathcal{Z}$:

$$\tilde{I}_r \leq \frac{\zeta^2}{2} \text{Tr} \left[\hat{Q}_r \left(\sum_j p_j \hat{D}_j \hat{\rho}^{(0)} \hat{D}_j \right) \right]. \quad (19)$$

Note that the obtained bound is quadratic in the cost parameter ζ .

In the case of type- \mathcal{Z}^\perp measurement operators, the condition $\text{Tr}[\hat{Q}_r \hat{\rho}^{(0)}] = 0$ combined with positive semidefiniteness of the operators \hat{Q}_r and $\hat{\rho}^{(0)}$ implies that $\hat{Q}_r \hat{\rho}^{(0)} = \hat{\rho}^{(0)} \hat{Q}_r = 0$. Therefore, a calculation of $\text{Tr}[\hat{Q}_r \hat{\rho}_j^{(1)}]$ using the SLD defined in equation (13) yields zero and the leading-order term in the power series expansion of the conditional probability $p_{r|j}(\zeta) = \text{Tr}[\hat{Q}_r \hat{\rho}_j(\zeta)]$ is quadratic in ζ . We will write it in the form

$$p_{r|j}(\zeta) \approx \zeta^2 \text{Tr}[\hat{Q}_r \hat{\rho}_j^{(2)}], \quad (20)$$

where

$$\hat{\rho}_j^{(2)} = \hat{\Pi} \hat{\rho}_j^{(2)} \hat{\Pi} \quad (21)$$

with $\hat{\rho}_j^{(2)}$ defined in equation (2) and $\hat{\Pi}$ being the projection onto the kernel of the zero-cost state $\hat{\rho}^{(0)}$. This follows from the fact that for type- \mathcal{Z}^\perp operators one has $\hat{\Pi} \hat{Q}_r \hat{\Pi} = \hat{Q}_r$.

As a result, the contribution to mutual information generated by a measurement outcome $r \in \mathcal{Z}^\perp$ can be written in the form:

$$I_r \approx \tilde{I}_r = \zeta^2 \sum_j p_j \text{Tr}[\hat{Q}_r \hat{\rho}_j^{(2)}] \log \frac{\text{Tr}[\hat{Q}_r \hat{\rho}_j^{(2)}]}{\sum_l p_l \text{Tr}[\hat{Q}_r \hat{\rho}_l^{(2)}]}. \quad (22)$$

Interestingly, although type- \mathcal{Z}^\perp operators generate events with vanishing probability in the limit $\zeta \rightarrow 0$, their contribution to mutual information is of the same order in ζ as from type- \mathcal{Z} operators. A similar behavior was observed for mutual information in a somewhat different context in [19].

Inserting equations (19) and (22) into (10), the leading-order expansion of mutual information in the cost parameter ζ is upper bounded by the following expression:

$$I \approx \tilde{I} \leq \zeta^2 \sum_{r \in \mathcal{Z}^\perp} \sum_j p_j \text{Tr}[\hat{Q}_r \hat{\rho}_j^{(2)}] \log \frac{\text{Tr}[\hat{Q}_r \hat{\rho}_j^{(2)}]}{\sum_l p_l \text{Tr}[\hat{Q}_r \hat{\rho}_l^{(2)}]} + \frac{\zeta^2}{2} \text{Tr} \left[\left(\sum_{r \in \mathcal{Z}} \hat{Q}_r \right) \left(\sum_j p_j \hat{D}_j \hat{\rho}^{(0)} \hat{D}_j \right) \right]. \quad (23)$$

The sum over \mathcal{Z} -type measurement operators appearing in the second line of the above formula can be expressed in terms of \mathcal{Z}^\perp -type operators as $\sum_{r \in \mathcal{Z}} \hat{Q}_r = \hat{\mathbb{I}} - \sum_{r \in \mathcal{Z}^\perp} \hat{Q}_r$. This allows us to write the derived bound solely in terms of \mathcal{Z}^\perp -type operators as

$$I \approx \tilde{I} \leq \zeta^2 \sum_{r \in \mathcal{Z}^\perp} \sum_j p_j \text{Tr}[\hat{Q}_r \hat{\rho}_j^{(2)}] \log \frac{\text{Tr}[\hat{Q}_r \hat{\rho}_j^{(2)}]}{\sum_l p_l \text{Tr}[\hat{Q}_r \hat{\rho}_l^{(2)}]} + \frac{\zeta^2}{2} \text{Tr} \left[\left(\hat{\mathbb{I}} - \sum_{r \in \mathcal{Z}^\perp} \hat{Q}_r \right) \left(\sum_j p_j \hat{D}_j \hat{\rho}_j^{(0)} \hat{D}_j \right) \right]. \quad (24)$$

As we will see in the following sections, the above expression, despite its seemingly complicated form, can help substantially in the analysis of specific communication scenarios in the low-cost limit.

4. Individual measurements on single-mode coherent states

The asymptotic analysis presented in section 3 can be applied to a variety of physical situations. In the following we will consider a scenario motivated by optical communication, when information is transmitted using an alphabet of coherent states of a single light mode [20, 21]

$$|\zeta\gamma_j\rangle = e^{-\zeta^2|\gamma_j|^2/2} \sum_{n=0}^{\infty} \frac{(\zeta\gamma_j)^n}{\sqrt{n!}} |n\rangle. \quad (25)$$

Here $|n\rangle$ denotes the n -photon Fock state. The complex amplitudes $\zeta\gamma_j$ of coherent states are written as products of complex coefficients γ_j and the nonnegative cost parameter ζ that can be viewed as the scaling factor for the entire ensemble represented as a constellation in the complex plane. Physically, it can describe the attenuation of an optical channel over which information transfer takes place. When $\zeta \rightarrow 0$, all the states converge to the vacuum state $|0\rangle$ which plays the role of the zero-cost state $\hat{\rho}^{(0)} = |0\rangle\langle 0|$ introduced in section 2. If individual coherent states $|\zeta\gamma_j\rangle$ are used with respective probabilities p_j , the average optical energy for the ensemble, measured in the photon number units, reads

$$\bar{n} = \zeta^2 \sum_j p_j |\gamma_j|^2 \quad (26)$$

and is proportional to the square of the cost parameter. As a simple example, m -ary phase shift keying (PSK) corresponds to taking $\gamma_j = e^{2\pi i j/m}$ and $p_j = 1/m$ with $j = 0, 1, \dots, m-1$. With this parametrization, all the input states have the optical energy equal to the average $\bar{n} = \zeta^2$.

In this section we consider a scenario when information is retrieved by measuring individually symbols in the form of single-mode coherent states defined in equation (25). It is straightforward to obtain explicit expressions for the two operators that enter the asymptotic expression for the mutual information given in equation (23). The SLD difference defined in equation (16) reads

$$\hat{D}_j = 2[(\gamma_j - \gamma_{\text{ens}})|1\rangle\langle 0| + (\gamma_j - \gamma_{\text{ens}})^*|0\rangle\langle 1|], \quad (27)$$

where $\gamma_{\text{ens}} = \sum_j p_j \gamma_j$, while

$$\hat{\rho}_j^{(2)} = |\gamma_j|^2 |1\rangle\langle 1|. \quad (28)$$

Consequently, it is sufficient to consider the action of measurement operators \hat{Q}_r only in the two-dimensional subspace spanned by the zero- and the one-photon Fock states $|0\rangle$ and $|1\rangle$.

4.1. Photon information efficiency bounds

Suppose first that the measurement is composed only of type- \mathcal{Z} operators. Inserting equation (27) into equation (19) and recalling that $\hat{\rho}^{(0)} = |0\rangle\langle 0|$ yields for $r \in \mathcal{Z}$:

$$\tilde{I}_r \leq 2\zeta^2 \langle 1|\hat{Q}_r|1\rangle \sum_j p_j |\gamma_j - \gamma_{\text{ens}}|^2. \quad (29)$$

The sum over j in the above formula gives the variance of the complex coefficients γ_j , which is upper bounded by the mean square $\sum_j p_j |\gamma_j|^2$. If all the measurement operators are type- \mathcal{Z} , one has $\sum_r \langle 1|\hat{Q}_r|1\rangle = 1$ which yields:

$$\tilde{I} \leq 2\zeta^2 \sum_j p_j |\gamma_j|^2 = 2\bar{n}, \quad (30)$$

where in the second step we have used the calculation of the ensemble average photon number carried out in equation (26). In optical communication, the ratio I/\bar{n} is known as the *photon information efficiency* (PIE) specifying the maximum amount of information that can be carried by one received photon. Therefore equation (30) can be viewed as an upper bound of 2 nats per photon on the PIE in the asymptotic limit $\bar{n} \rightarrow 0$, which holds for measurements composed from type- \mathcal{Z} operators \hat{Q}_r satisfying the condition $\langle 0|\hat{Q}_r|0\rangle > 0$.

Interestingly, the asymptotic bound of 2 nats per photon on the PIE in the limit $\bar{n} \rightarrow 0$ appears also in a scenario when information is encoded in a single quadrature of the optical mode measured by shot-noise limited homodyne detection [11]. In this case the so-called Shannon–Hartley limit for the mutual information reads [22–26]

$$I_{\text{SH}} = \frac{1}{2} \log(1 + 4\bar{n}) \quad (31)$$

and is attained by using an input ensemble of coherent states with the Gaussian distribution for the measured quadrature and the conjugate quadrature set to zero. A semi-qualitative connection between this scenario and the bound derived in equation (30) can be made as follows. As it is well known, decomposing POVM elements into positive rank-1 operators can make mutual information only equal or greater [27, 28]. Type- \mathcal{Z} and rank-1 POVM elements \hat{Q}_r with non-zero matrix elements $\langle 1|\hat{Q}_r|1\rangle$ which non-trivially contribute to equation (30) must detect coherence between zero- and one-photon Fock states and consequently are sensitive to the phase of the input coherent state. Homodyne detection is an example of such a phase-sensitive measurement, although in the case of the homodyne POVM the classification criterion based on whether $\text{Tr}[\hat{Q}_r \hat{\rho}^{(0)}] = 0$ or not used in section 3 becomes problematic. This is because the outcome of the homodyne measurement has the continuous form corresponding to the field quadrature which can take any real value and the conditional probabilities for homodyne POVM elements on the zero-cost vacuum state become arbitrarily small with the increasing absolute value of the quadrature. More generally, the low-cost limit of PIE attainable with individual phase-sensitive measurements coincides with that derived from the Shannon–Hartley theorem also when the input coherent states are transmitted over a phase-invariant Gaussian channel with excess noise, mapping them onto displaced thermal states. This case is discussed in appendix A.

Let us now turn our attention to type- \mathcal{Z}^\perp measurement operators \hat{Q}_r . The condition $\text{Tr}[\hat{Q}_r \hat{\rho}^{(0)}] = 0$ defining type- \mathcal{Z}^\perp operators implies that in the zero- and one-photon subspace \hat{Q}_r must be proportional to a projection $|1\rangle\langle 1|$ onto the one-photon state. Therefore in optical scenarios type- \mathcal{Z}^\perp operators correspond to direct photon counting. Inserting equation (28) into (22) yields the contribution to mutual information from a measurement outcome $r \in \mathcal{Z}^\perp$ in the form

$$\tilde{I}_r = \zeta^2 \langle 1|\hat{Q}_r|1\rangle \sum_j p_j |\gamma_j|^2 \log \frac{|\gamma_j|^2}{\sum_l p_l |\gamma_l|^2}. \quad (32)$$

A frequently encountered constraint in optical communication systems is an upper limit on the peak-to-average power ratio \mathcal{P} of the transmitted signal. In our case \mathcal{P} is given by the ratio of the maximum mean photon number carried by one symbol to the average photon number in the input ensemble:

$$\mathcal{P} = \frac{\max_j |\gamma_j|^2}{\sum_l p_l |\gamma_l|^2}. \quad (33)$$

If all the measurement operators are type- \mathcal{Z}^\perp , equation (32) summed over r and combined with equation (33) provides an upper bound on mutual information in the form:

$$\tilde{I} \leq \zeta^2 \sum_j p_j |\gamma_j|^2 \log \mathcal{P} \leq \bar{n} \log \mathcal{P}, \quad (34)$$

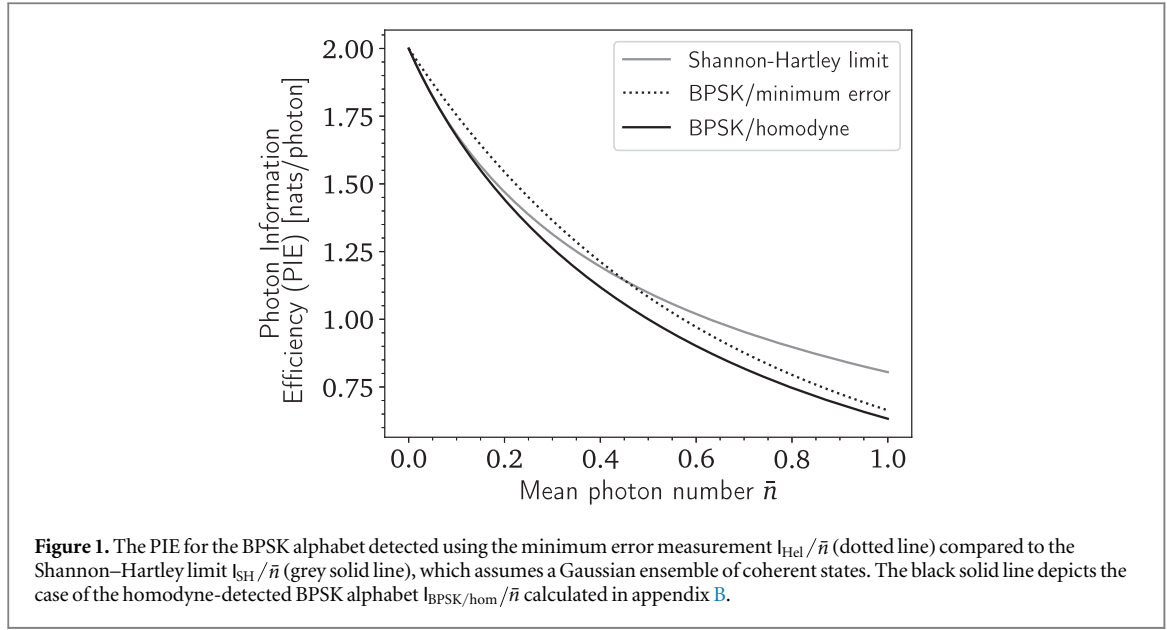
where in the second step we have used equation (26). The above inequality implies that when photon counting detection is used, the PIE is upper bounded by the logarithm of the peak-to-average power ratio \mathcal{P} . It is easy to verify that the maximum value $\log \mathcal{P}$ can be achieved using a two-symbol ensemble consisting of a coherent state $|\sqrt{\mathcal{P}\bar{n}}\rangle$ sent with a probability $1/\mathcal{P}$ and the vacuum state $|0\rangle$ sent with the probability $1 - 1/\mathcal{P}$, detected by Geiger-mode photon counting which discriminates between zero photons and at least one photon. Such a strategy is customarily used in deep-space optical communication, where photon information efficiency is of primary importance [29, 30]. The practical implementation is based on the pulse position modulation (PPM) format which uses frames composed of an integer number of \mathcal{P} temporal bins. Each frame contains one pulse with the mean photon number $\mathcal{P}\bar{n}$ located in one of the bins that are otherwise left empty.

Note that equation (32) implies that for PSK alphabets, when all γ_j are equal in their magnitude, the contribution to mutual information from type- \mathcal{Z}^\perp operators is zero, $\tilde{I}_r = 0$ when $r \in \mathcal{Z}^\perp$. This is consistent with the simple fact that direct photon counting cannot reveal information encoded in the phase of the electromagnetic field. Consequently, a general detection strategy combining type- \mathcal{Z} and type- \mathcal{Z}^\perp operators for a PSK ensemble has the PIE upper bounded by 2 nats per photon in the limit $\bar{n} \rightarrow 0$.

4.2. Individual measurements on BPSK symbols

Let us now discuss the case of individual measurements on the binary phase shift keying (BPSK) alphabet of two coherent states $|\zeta\rangle$ and $|\!-\!\zeta\rangle$ with opposite phases and the same mean photon number $\bar{n} = \zeta^2$. The maximum mutual information attainable in this scenario can be written in a closed analytical form for a finite \bar{n} . We will see that the BPSK alphabet is sufficient to saturate the 2 nats per photon bound on the PIE in the low-cost limit. The exact expression for the mutual information attainable for a given \bar{n} will be used in section 5 as a benchmark when discussing superadditive measurements on multiple BPSK symbols.

It has been shown [31, 32] that when a binary ensemble of equiprobable pure states is used for classical information transmission and only individual detection is permitted, mutual information is maximized by the



minimum-error measurement described by Helstrom [13]. The symmetric probability of a wrong identification of the input BPSK state can be obtained from the Helstrom bound

$$\varepsilon_{\text{Hel}} = \frac{1}{2}(1 - \sqrt{1 - |\langle -\zeta|\zeta \rangle|^2}) = \frac{1}{2}(1 - \sqrt{1 - e^{-4\bar{n}}}), \quad (35)$$

which yields mutual information expressed in nats:

$$I_{\text{Hel}} = \log 2 - H(\varepsilon_{\text{Hel}}). \quad (36)$$

Here

$$H(x) = -x \log x - (1 - x) \log(1 - x) \quad (37)$$

is the entropy of a binary random variable specified in nats. The resulting exact expression for the PIE attainable with individual measurements on the BPSK alphabet, equal to I_{Hel}/\bar{n} , is depicted in figure 1 as a function of \bar{n} . In the limit $\bar{n} \rightarrow 0$ the PIE reaches 2 nats. It is worth noting that for $\bar{n} \lesssim 0.4$ the obtained result exceeds the PIE I_{SH}/\bar{n} obtained from the Shannon–Hartley expression given by equation (31), also plotted in figure 1. This highlights the specificity of the Shannon–Hartley limit, used as a canonical reference in optical communication, to the quadrature measurement assumed in its derivation. The lower value of I_{SH}/\bar{n} compared to I_{Hel}/\bar{n} for $\bar{n} \ll 1$ can be verified by inspecting the second-to-leading terms in the asymptotic expansions of the respective expressions given in equations (31) and (36):

$$I_{\text{SH}} \approx 2\bar{n} - 4\bar{n}^2, \quad I_{\text{Hel}} \approx 2\bar{n} - \frac{8}{3}\bar{n}^2. \quad (38)$$

As also shown in figure 1, the PIE of 2 nats can be attained in the limit $\bar{n} \rightarrow 0$ with a homodyne measurement on the BPSK alphabet, provided that the information is retrieved from the measured continuous quadrature value. For completeness, the mathematical description of this scenario is presented in appendix B. Of course, in this case the PIE does not exceed I_{SH}/\bar{n} for $\bar{n} > 0$.

The physical realization of the minimum error measurement for the BPSK alphabet is known to be the Dolinar receiver based on photon counting of the displaced input signal with a fast feedback loop [33]. Unfortunately, due to its complexity, realizations of the Dolinar receiver have so far been confined to the proof-of-principle stage [34]. In search for technically less challenging implementations, alternative nearly-optimal measurements have been investigated, including a feed-forward scheme based on signal splitting and optical delay lines [35]. A considerable amount of work has been devoted also to displacement receivers without feedback or feed-forward mechanisms [36–42] which show a pretty good performance. Additionally, a non-demolition measurement based on an atom-light interaction that comes close to the Helstrom bound has been recently proposed [43]. Furthermore, derivatives of the Dolinar receiver have been demonstrated for alphabets consisting of more than two coherent states like for example commonly considered quaternary phase shift keying or higher PSK formats [44–50]. In addition to the minimum-error criterion, one can consider other strategies to distinguish between non-orthogonal quantum states, such as unambiguous state discrimination [51–53]. Time-resolved photon counting of the PPM format mentioned in section 4.1 can be viewed as a realization of the unambiguous strategy for quantum states of light prepared within PPM frames: registering a

photocount in a given time bin identifies unambiguously the location of the pulse, i.e. the input state, whereas absence of any photocount over the entire frame yields an inconclusive result.

5. Collective measurements on BPSK words

We will now use the upper bound on the mutual information expanded in the cost parameter as derived in section 3 to identify collective measurement schemes on multiple BPSK symbols that beat the performance of individual detection. While the designs for collective schemes will be motivated by the analysis of the mutual information in the asymptotic limit of the vanishing cost, the actual figure for the mutual information, or equivalently the photon information efficiency, attainable using the designed measurement will be calculated exactly for a given finite \bar{n} and compared against the individual detection benchmark given in equation (36).

When collective measurements can be performed on a sequence of individual quantum systems, for example a train of single-mode light pulses, one needs to consider an input ensemble in the form of *words* composed from the alphabet of elementary symbols, such as coherent states introduced in equation (25). The input probabilities are then defined for entire words. In the remainder of the paper, we will restrict our attention to words constructed from the BPSK alphabet of coherent states $|\zeta\rangle$ and $|\zeta^*\rangle$. Words of length M can be then conveniently labelled with M -bit strings $\mathbf{j} = j_1 j_2 \dots j_M$. The input states are pure, $\hat{\rho}_{\mathbf{j}} = |\psi_{\mathbf{j}}\rangle \langle \psi_{\mathbf{j}}|$, and have the form of M -mode coherent states,

$$|\psi_{\mathbf{j}}\rangle = |(-1)^{j_1} \zeta\rangle |(-1)^{j_2} \zeta\rangle \dots |(-1)^{j_M} \zeta\rangle. \quad (39)$$

The label \mathbf{j} needs now to be used as an ensemble index in lieu of j when applying general results derived in section 3. For a fair comparison with the individual measurement scenario, the cost should be quantified with the mean photon number *per elementary symbol*, given in the present case by $\bar{n} = \zeta^2$. The entire optical energy of a BPSK word of length M is $M\zeta^2 = M\bar{n}$.

Our goal will be to identify the probability distribution $p_{\mathbf{j}}$ for input words $|\psi_{\mathbf{j}}\rangle$ and the few-symbol collective measurement strategy that maximizes mutual information in the low-cost limit. Let us first specialize equation (24) to the scenario considered here. It will be convenient to denote a normalized superposition state of one photon in M modes as:

$$|1_{\mathbf{j}}\rangle = \frac{1}{\sqrt{M}} [(-1)^{j_1} |1\rangle |0\rangle \dots |0\rangle + (-1)^{j_2} |0\rangle |1\rangle \dots |0\rangle + \dots + (-1)^{j_M} |0\rangle |0\rangle \dots |1\rangle] \quad (40)$$

and use $|v\rangle = |0\rangle |0\rangle \dots |0\rangle$ for the M -mode vacuum state. With this notation, the SLD difference for the state $\hat{\rho}_{\mathbf{j}}$ reads

$$\hat{D}_{\mathbf{j}} = 2\sqrt{M} (|1_{\mathbf{j}}\rangle \langle v| + |v\rangle \langle 1_{\mathbf{j}}|) \quad (41)$$

and respectively

$$\hat{\rho}_{\mathbf{j}}^{(2)} = M |1_{\mathbf{j}}\rangle \langle 1_{\mathbf{j}}|. \quad (42)$$

After inserting these expressions into equation (24) it is easy to notice that pairs of terms corresponding to \mathbf{j} and its bitwise negation $\neg\mathbf{j}$ can be combined together because the corresponding one-photon states $|1_{\mathbf{j}}\rangle$ and $|1_{\neg\mathbf{j}}\rangle$ differ only by an irrelevant global minus sign. We will select one of the \mathbf{j} 's as a representative for each pair and use the equivalence class $[\mathbf{j}]$ as an index for the two-fold reduced sum over the input ensemble. The result reads

$$\tilde{I} \leq 2M\zeta^2 + M\zeta^2 \sum_{r \in \mathcal{Z}^{\perp}} \sum_{[\mathbf{j}]} p_{[\mathbf{j}]} \langle 1_{[\mathbf{j}]} | \hat{Q}_r | 1_{[\mathbf{j}]} \rangle \left(\log \frac{\langle 1_{[\mathbf{j}]} | \hat{Q}_r | 1_{[\mathbf{j}]} \rangle}{\sum_{[\mathbf{l}]} p_{[\mathbf{l}]} \langle 1_{[\mathbf{l}]} | \hat{Q}_r | 1_{[\mathbf{l}]} \rangle} - 2 \right), \quad (43)$$

where $p_{[\mathbf{j}]} = p_{\mathbf{j}} + p_{\neg\mathbf{j}}$ is the total probability of using either word from a given pair. The above expression can be equivalently written as

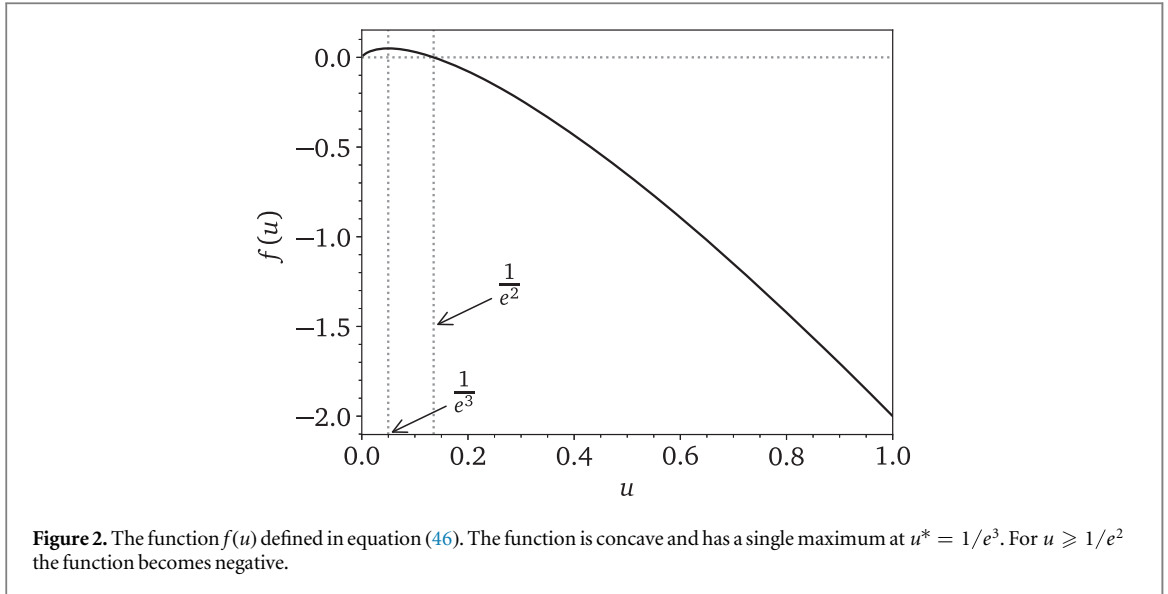
$$\tilde{I} \leq 2M\zeta^2 + M\zeta^2 \sum_{r \in \mathcal{Z}^{\perp}} \sum_{[\mathbf{j}]} f(p_{[\mathbf{j}]}) \langle 1_{[\mathbf{j}]} | \hat{Q}_r | 1_{[\mathbf{j}]} \rangle \quad (44)$$

$$+ M\zeta^2 \sum_{r \in \mathcal{Z}^{\perp}} \sum_{[\mathbf{j}]} p_{[\mathbf{j}]} \langle 1_{[\mathbf{j}]} | \hat{Q}_r | 1_{[\mathbf{j}]} \rangle \log \frac{p_{[\mathbf{j}]} \langle 1_{[\mathbf{j}]} | \hat{Q}_r | 1_{[\mathbf{j}]} \rangle}{\sum_{[\mathbf{l}]} p_{[\mathbf{l}]} \langle 1_{[\mathbf{l}]} | \hat{Q}_r | 1_{[\mathbf{l}]} \rangle}, \quad (45)$$

where

$$f(u) = u \log \frac{1}{u} - 2u. \quad (46)$$

Since for any \mathbf{j} and r the argument of the logarithm in equation (44) does not exceeds one, the second line of that expression containing logarithms is nonpositive therefore by neglecting it one arrives at a weaker upper bound of the form:



$$\tilde{I} \leq 2M\zeta^2 + M\zeta^2 \sum_{[j]} f(p_{[j]}) \langle l_{[j]} | \left(\sum_{r \in \mathcal{Z}^\perp} \hat{Q}_r \right) | l_{[j]} \rangle. \quad (47)$$

The function $f(u)$ defined in equation (46) will play an important role in further analysis. Its graph is shown in figure 2. It is easy to verify by an elementary calculation that on the interval $0 \leq u \leq 1$ the function $f(u)$ is concave, has a single maximum at $u^* = 1/e^3 \approx 0.0498$ equal to $f(u^*) = u^*$ and becomes negative for arguments $u > 1/e^2$.

5.1. Two-symbol measurements

In the case of two-symbol measurements, $M = 2$, we have only two non-equivalent one-photon states that can be compactly denoted as

$$|+\rangle = |l_{00}\rangle = -|l_{11}\rangle = \frac{1}{\sqrt{2}}(|1\rangle|0\rangle + |0\rangle|1\rangle), \quad (48)$$

$$|-\rangle = |l_{01}\rangle = -|l_{10}\rangle = \frac{1}{\sqrt{2}}(|1\rangle|0\rangle - |0\rangle|1\rangle). \quad (49)$$

We will also use $+$ and $-$ symbols to label the two respective equivalence classes $+\equiv[00] = [11]$ and $+\equiv[01] = [10]$ of the input words with corresponding probabilities $p_+ = p_{00} + p_{11}$ and $p_- = p_{10} + p_{01}$. Equation (47) can be then simplified to the form

$$\tilde{I} \leq 2\bar{n}[2 + Q^+ f(p_+) + Q^- f(p_-)], \quad (50)$$

where we have denoted

$$Q^+ = \sum_{r \in \mathcal{Z}^\perp} \langle + | \hat{Q}_r | + \rangle, \quad Q^- = \sum_{r \in \mathcal{Z}^\perp} \langle - | \hat{Q}_r | - \rangle. \quad (51)$$

Because Q^\pm are the overall probabilities of obtaining an $r \in \mathcal{Z}^\perp$ outcome on states $|\pm\rangle$, they are constrained by $0 \leq Q^\pm \leq 1$. The goal now is to maximize the right hand side of equation (50). Given the properties of the function $f(u)$ discussed earlier and taking into account that $p_+ + p_- = 1$, the right hand side of equation (50) is optimized by selecting $p_- = u^* = 1/e^3$, $Q^- = 1$, and $Q^+ = 0$. An equivalent solution has exchanged labels $+$ and $-$. This finally yields:

$$\tilde{I} \leq 2\bar{n}[2 + f(u^*)] = 2\bar{n}(2 + 1/e^3). \quad (52)$$

Given that the mean photon number in an input word is $2\bar{n}$, the asymptotic value of the PIE is upper bounded by $\tilde{I}/(2\bar{n}) \leq 2 + 1/e^3$, which presents a relative increase of $1/(2e^3) \approx 2.49\%$ compared to the individual measurement scenario. Note that the value derived in equation (52) maximizes also the full expression given in equation (44), as for the chosen p_\pm and Q_\pm the second logarithmic term in equation (44) vanishes.

The optimum identified above can be translated into an optical measurement scheme that asymptotically saturates the bound derived in equation (52). The condition $Q^- = 1$ implies that input states $|\psi_{01}\rangle$ and $|\psi_{10}\rangle$ containing the $|-\rangle$ component should be detected by type- \mathcal{Z}^\perp measurement operators, which as discussed in section 4 correspond in the optical domain to single photon detection. Since photon detection is phase-insensitive, the optimum $p_- = p_{01} + p_{10}$ specifies only the overall probability of using the two input states $|\psi_{01}\rangle$

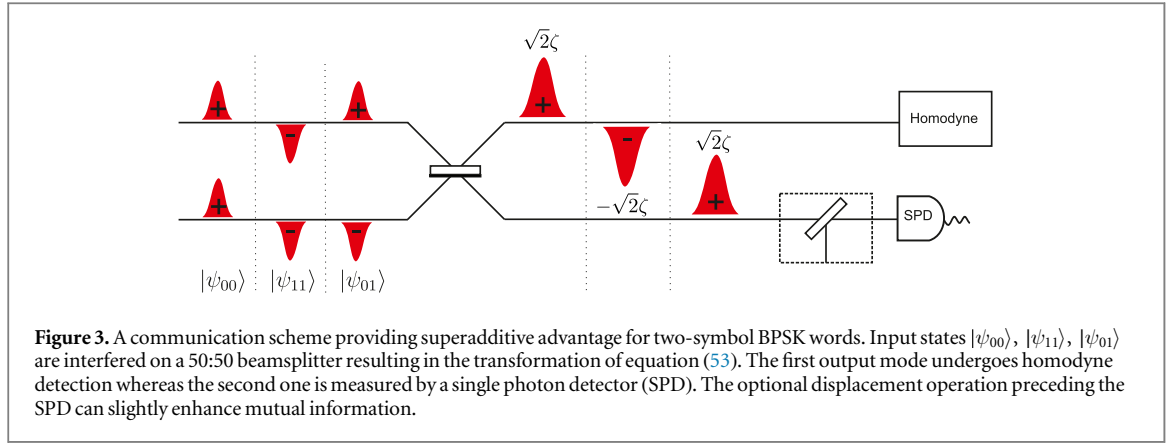


Figure 3. A communication scheme providing superadditive advantage for two-symbol BPSK words. Input states $|\psi_{00}\rangle$, $|\psi_{11}\rangle$, $|\psi_{01}\rangle$ are interfered on a 50:50 beamsplitter resulting in the transformation of equation (53). The first output mode undergoes homodyne detection whereas the second one is measured by a single photon detector (SPD). The optional displacement operation preceding the SPD can slightly enhance mutual information.

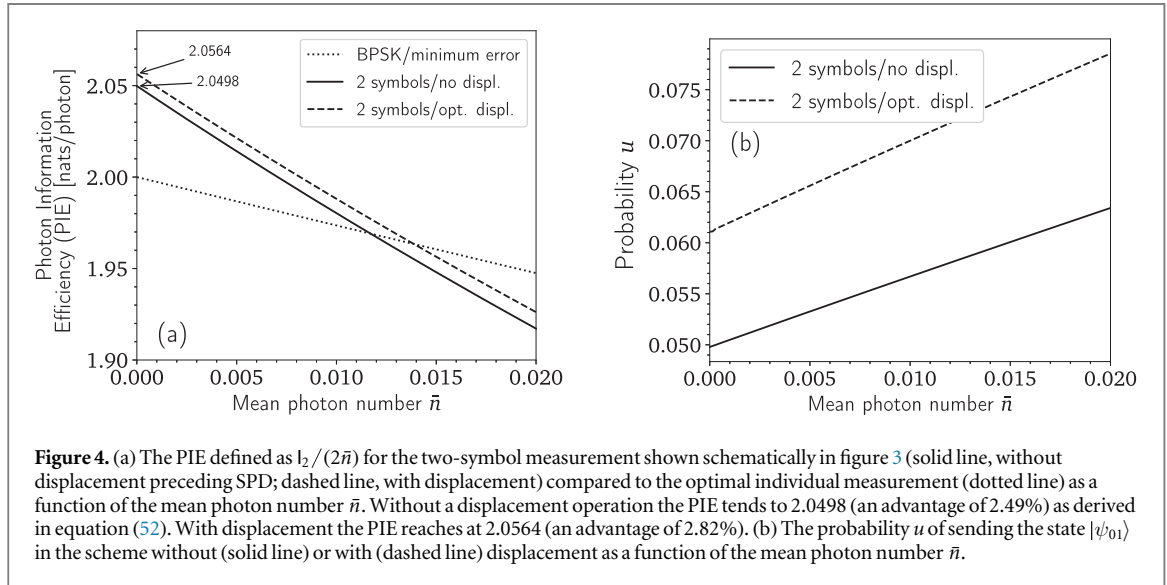


Figure 4. (a) The PIE defined as $I_2/(2\bar{n})$ for the two-symbol measurement shown schematically in figure 3 (solid line, without displacement preceding SPD; dashed line, with displacement) compared to the optimal individual measurement (dotted line) as a function of the mean photon number \bar{n} . Without a displacement operation the PIE tends to 2.0498 (an advantage of 2.49%) as derived in equation (52). With displacement the PIE reaches at 2.0564 (an advantage of 2.82%). (b) The probability u of sending the state $|\psi_{01}\rangle$ in the scheme without (solid line) or with (dashed line) displacement as a function of the mean photon number \bar{n} .

and $|\psi_{10}\rangle$. For concreteness, we will take only $|\psi_{01}\rangle$. On the other hand, the requirement $Q^+ = 0$ means that input states $|\psi_{00}\rangle$ and $|\psi_{11}\rangle$ containing the one-photon component $|+\rangle$ are measured only by type- \mathcal{L} operators, which corresponds to phase-sensitive detection. We have seen in section 4 that such a measurement can be realized by the Dolinar receiver or homodyning. In order to maximize the contribution to mutual information, one needs to use the states $|\psi_{00}\rangle$ and $|\psi_{11}\rangle$ with the same input probabilities $p_{00} = p_{11}$.

A straightforward way to subject the input states to the required type of detection is to interfere the two component coherent states in each two-symbol word on a balanced 50/50 beam splitter, as shown in figure 3. This results in the transformation

$$\begin{aligned} |\psi_{00}\rangle &\rightarrow |\sqrt{2}\zeta\rangle|0\rangle, \\ |\psi_{11}\rangle &\rightarrow |-\sqrt{2}\zeta\rangle|0\rangle, \\ |\psi_{01}\rangle &\rightarrow |0\rangle|\sqrt{2}\zeta\rangle, \end{aligned} \quad (53)$$

where the two kets on the right hand side refer to the two output ports of the beam splitter. Note that when a Dolinar receiver is used as a phase-sensitive measurement, this reproduces the two-symbol measurement strategy described by Guha [10]. In the following, we consider homodyne detection as a more standard phase-sensitive measurement technique. For a finite $\bar{n} = \zeta^2$ we have optimized numerically the PIE $I_2/(2\bar{n})$ given by the ratio of the actual mutual information I_2 calculated using equation (B6) in appendix B and the mean photon number over input probabilities parameterized with $p_{01} + p_{10} = u$, $p_{00} = p_{11} = (1 - u)/2$ in a scenario when a combination of homodyne detection and single photon detection (SPD) is implemented on the two output beam splitter ports. The results are shown in figure 4. It is seen that in the limit $\bar{n} \rightarrow 0$ the PIE indeed approaches the value derived in equation (52) which is also shown analytically in appendix B. Superadditivity of accessible information for the BPSK alphabet, when the amount of transmitted information exceeds that attainable with individual measurements, occurs for $\bar{n} < 0.0117$.

The PIE attainable for two-symbol measurements can be further improved by introducing a displacement operation before the single photon detector (SPD), shown within a dashed box in figure 3. This operation adds a

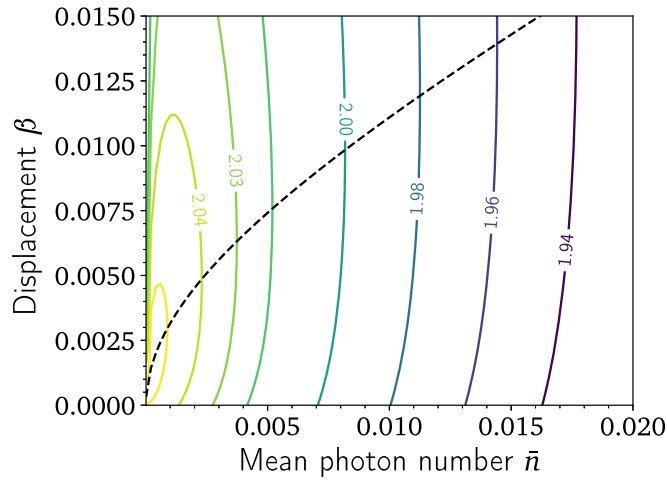


Figure 5. A contour plot of the PIE $I/(2\bar{n})$ for the two-symbol measurement depicted in figure 3 as a function of the mean photon number \bar{n} and the displacement β . The dashed line indicates the optimal displacement value for a given mean photon number \bar{n} .

coherent amplitude β to the field detected by the SPD. Such an arrangement is known as the generalized Kennedy receiver [38, 40–42]. As seen in figure 4, optimizing β for a given \bar{n} yields a slightly higher value of PIE, approaching 2.0564 in the limit $\bar{n} \rightarrow 0$, which corresponds to 2.82% enhancement relative to individual measurements. Noticeably, the displacement value β tends to zero with diminishing \bar{n} , as seen in figure 5. This result indicates that the asymptotic analysis of mutual information in the low-cost limit carried out in section 3 may yield a different bound if explicit dependence of the measurement operators on the cost parameter is allowed. A similar effect occurs also in the analysis of superadditivity of accessible information in the case of a collective measurement on qubit pairs [9].

5.2. Three-symbol measurements

In the case of three-symbol BPSK words, the one-photon sector relevant to the evaluation of mutual information in the low-cost limit is spanned by four states:

$$\begin{aligned}
 |++\rangle &= |1_{000}\rangle = -|1_{111}\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle), \\
 |+-\rangle &= |1_{001}\rangle = -|1_{110}\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle - |001\rangle), \\
 |-+\rangle &= |1_{010}\rangle = -|1_{101}\rangle = \frac{1}{\sqrt{3}}(|100\rangle - |010\rangle + |001\rangle), \\
 |--\rangle &= |1_{011}\rangle = -|1_{100}\rangle = \frac{1}{\sqrt{3}}(|100\rangle - |010\rangle - |001\rangle).
 \end{aligned} \tag{54}$$

An essential difference with the two-symbol case is that the four single-photon states listed above are mutually nonorthogonal. This makes it problematic to use equation (47) for optimization as the matrix elements $\langle 1_{[j]} | (\sum_{r \in \mathcal{Z}^1} \hat{Q}_r) | 1_{[j]} \rangle$ cannot be set independently for each one photon state $|1_{[j]}\rangle$. Consequently, one needs to revert back to equation (43) for further analysis. We will denote the four equivalence classes with indices $++ \equiv [000] = [111]$, $+- \equiv [001] = [110]$, $-+ \equiv [010] = [101]$, and $-- \equiv [011] = [100]$.

We have performed a numerical search for input probabilities $p_{[j]}$ and type- \mathcal{Z}^1 POVM elements of rank-1 form in the one-photon sector that maximize the right hand side of equation (43). Up to a permutation of the modes, the numerical results suggest an optimum characterized by a high degree of symmetry with

$$p_{++} = 1 - 2\nu, \quad p_{+-} = p_{-+} = \nu, \quad p_{--} = 0 \tag{55}$$

and two type- \mathcal{Z}^1 projective operators $\hat{Q}_r = |q_r\rangle \langle q_r|$, $r = 1, 2$, which are given in the orthonormal basis $|100\rangle, |010\rangle, |001\rangle$ by

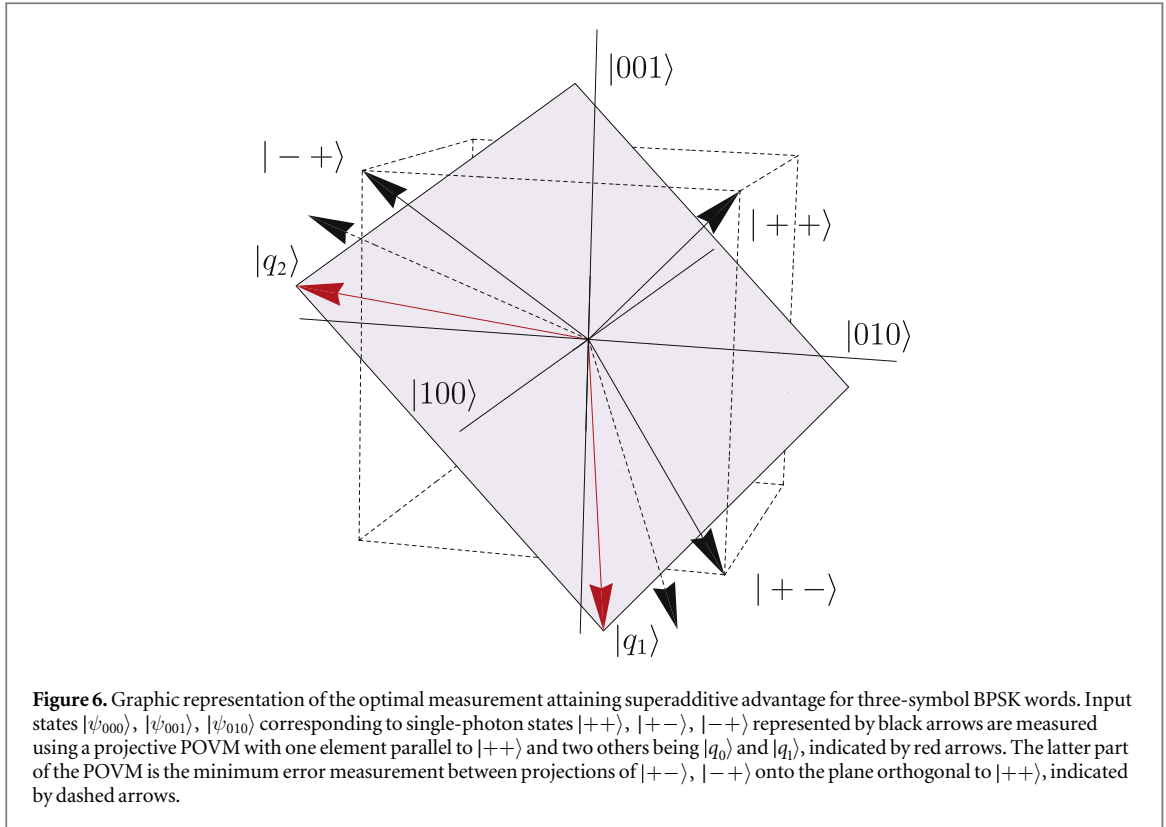


Figure 6. Graphic representation of the optimal measurement attaining superadditive advantage for three-symbol BPSK words. Input states $|\psi_{000}\rangle$, $|\psi_{001}\rangle$, $|\psi_{010}\rangle$ corresponding to single-photon states $|++\rangle$, $|+-\rangle$, $|--\rangle$ represented by black arrows are measured using a projective POVM with one element parallel to $|++\rangle$ and two others being $|q_0\rangle$ and $|q_1\rangle$, indicated by red arrows. The latter part of the POVM is the minimum error measurement between projections of $|+-\rangle$, $|--\rangle$ onto the plane orthogonal to $|++\rangle$, indicated by dashed arrows.

$$|q_1\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) \\ -\frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) \end{pmatrix}, \quad |q_2\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) \\ \frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) \end{pmatrix}. \quad (56)$$

It is straightforward to verify that $\langle q_1|++\rangle = \langle q_2|++\rangle = 0$. The geometry of the measurement is visualized in figure 6 with a three-dimensional diagram using the fact that the relevant state vectors have real components. The four one-photon states specified in equation (54) are located in vertices of a cube centered at the origin of the coordinate system. The two mutually orthogonal states $|q_1\rangle$ and $|q_2\rangle$ corresponding to projective type- \mathcal{Z}^\perp operators lie in the plane perpendicular to $|++\rangle$. This means that the vector $|++\rangle$ defines a subspace on which a type- \mathcal{Z} measurement needs to be carried out. The vectors $|q_1\rangle$ and $|q_2\rangle$ can be viewed as an implementation of a minimum-error measurement for inputs $|+-\rangle$ and $|--\rangle$ projected onto the plane perpendicular to $|++\rangle$.

Inserting equation (55) and type- \mathcal{Z}^\perp operators specified in equation (56) into (44) yields

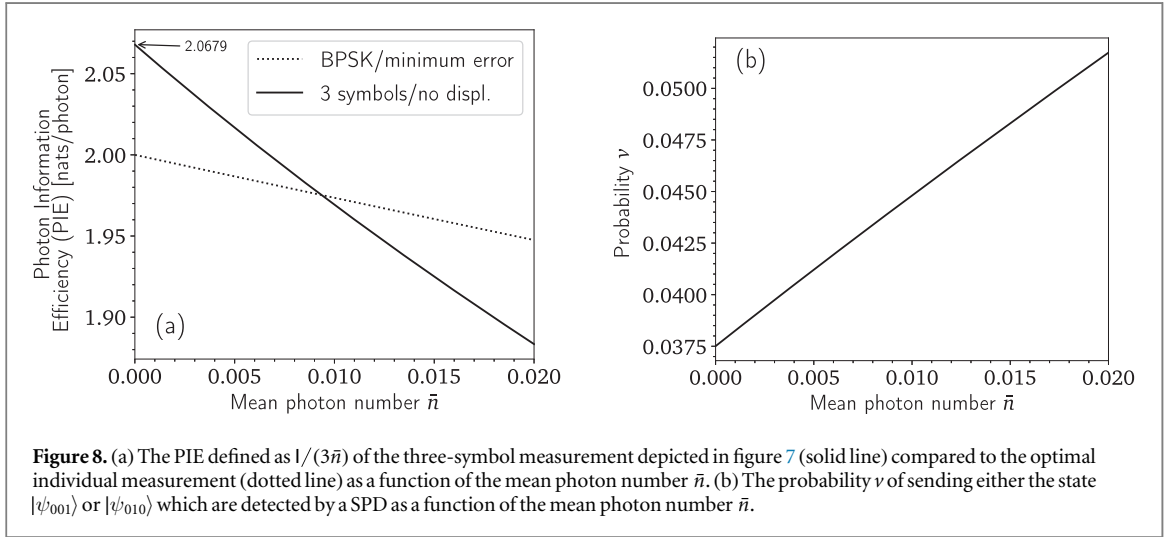
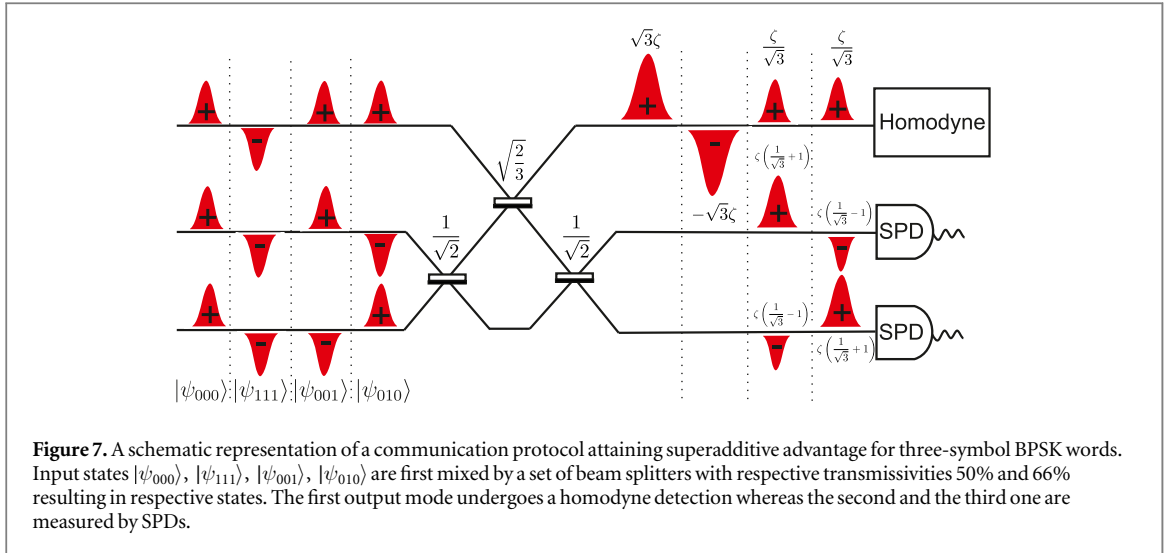
$$\tilde{I}_3 \leq 6\bar{n} + \frac{6\bar{n}}{9} \{4\nu[\sqrt{3} \log(2 + \sqrt{3}) - 4\log 2 - \nu] + 8f(\nu)\}. \quad (57)$$

The right hand side reaches maximum at $\nu^* \approx 0.0375$ resulting in a PIE bound $\tilde{I}_3/(3\bar{n}) \leq 2.0679$, which represents relative 3.40% improvement compared to individual measurements.

Analogously to the two-symbol case, the above solution suggests an optical implementation shown in figure 7. A single three-port linear optical circuit is constructed such that photons prepared in one of three orthogonal superposition states $|++\rangle$, $|q_1\rangle$, and $|q_2\rangle$ are routed to different output ports. Homodyne detection is implemented on the port corresponding to the $|++\rangle$ input, while two other ports are monitored by SPDs. The input probabilities are parameterized as $p_{000} = p_{111} = 1/2 - \nu$ and $p_{001} = p_{010} = \nu$. Maximization over ν for a given $\bar{n} = \zeta^2$ yields the actual photon information efficiency $I_3/(3\bar{n})$ shown in figure 8(a) with the optimal value of ν depicted in figure 8(b). It is seen that the PIE and the probability ν attain respectively the 3.40% enhancement and the optimal value implied by equation (57). in the asymptotic limit $\bar{n} \rightarrow 0$.

5.3. Orthogonal words

Although for BPSK words longer than two symbols full optimization of the asymptotic expression for mutual information requires resorting to numerical means, some interesting observations can be made when the ensemble of input words is *a priori* restricted to a subset for which all the one-photon states defined in equation (40) are mutually orthogonal, apart from pairs that differ only by a global minus sign. Such subsets can be systematically constructed using Hadamard matrices, which are real orthogonal matrices with entries ± 1



defined in dimensions that are integer powers of 2 and conjectured to exist for any dimension that is a multiple of 4 [54]. The orthogonal words are chosen such that sequences of signs for individual coherent states specified in equation (39) are given by rows of a Hadamard matrix, up to a global phase flip. The orthogonality property of Hadamard matrices mean that for such a reduced subset one-photon states $|1_j\rangle$ and $|1_{j'}\rangle$ defined in equation (40) are mutually orthogonal when j and j' correspond to different rows of a Hadamard matrix.

The orthogonality of the states $|1_{[j]}\rangle$ implies that the value of each of the matrix elements $Q^{[j]} = \langle 1_{[j]} | (\sum_{r \in \mathcal{R}} \hat{Q}_r) | 1_{[j]} \rangle$ in equation (47) can be set independently within physical constraints $0 \leq Q^{[j]} \leq 1$. The obvious way to maximize equation (47) is then to take $Q^{[j]} = 1$ if $f(p_{[j]}) > 0$ and $Q^{[j]} = 0$ otherwise. Recalling that $f(t)$ is positive for arguments $0 < t < 1/e^2$, this reduces equation (47) to

$$\tilde{I} \leq 2M\zeta^2 + M\zeta^2 \sum_{[j]: 0 < p_{[j]} < 1/e^2} f(p_{[j]}). \quad (58)$$

Importantly, by taking measurement operators \hat{Q}_r in the form of projections onto one-photon states $|1_{[j]}\rangle \langle 1_{[j]}|$, the second line of equation (44) can be made identically equal to zero, which implies that the bound (58) is equivalent to equation (43).

Suppose now that exactly M' indices $[j]$ satisfy the condition $0 < p_{[j]} < 1/e^2$. The concavity of $f(t)$ on the interval $0 < t < 1/e^2$ implies that

$$\sum_{[j]: 0 < p_{[j]} < 1/e^2} f(p_{[j]}) \leq M' f\left(\frac{1}{M'} \sum_{[j]: 0 < p_{[j]} < 1/e^2} p_{[j]}\right). \quad (59)$$

This means that it is always beneficial to make all the probabilities $p_{[j]}$ from the interval $0 < p_{[j]} < 1/e^2$ identical. Consequently, the sum appearing on the right hand side of equation (58) is maximized by an

expression of the form $M'f(t)$ under constraints $0 < t < 1/e^2$ and $M't \leq 1$ to guarantee that the sum of the probabilities does not exceed one.

In the case of orthogonal words based on the Hadamard construction one has $M' \leq M$. In this case a straightforward calculation shows that the mutual information is maximized by one of two strategies. The first one uses $M' = M$ and $t = 1/M$. The resulting bound on the PIE has the simple form $\log M$. This value can be achieved by feeding Hadamard words into a linear optical circuit which maps them onto the PPM format subsequently detected by single photon detection [10, 12]. The second strategy is to use $M' = M - 1$ Hadamard sequences with probabilities equal to $t^* = 1/e^3$ which maximize individual terms in equation (58). The residual probability $1 - (M - 1)t^*$ should be divided equally between the remaining Hadamard sequence and its phase flipped version that need to be detected by a type- \mathcal{Z} measurement. The resulting PIE is $2 + (M - 1)/e^3$. The first strategy is optimal for $M \geq 16$, while the second one offers higher PIE for $M = 2, 4, 8, 12$. Its special case for $M = 2$ is the measurement described in section 5.1. These observations are consistent with the results presented in [55]. It should be kept in mind that the limiting value of PIE is attained when $M\bar{n} \ll 1$ and terms of the order $(M\bar{n})^2$ and higher can be neglected when calculating the photocount probability. When selecting the communication protocol for a given \bar{n} it may be optimal to choose a finite M [56].

6. Conclusions

We presented a systematic expansion of mutual information for a given ensemble of quantum states and a fixed measurement in the cost parameter characterizing the input ensemble, assuming that with vanishing cost all states converge to one zero-cost state. In the case of optical communication with coherent states, the cost is naturally quantified in terms of the average photon number and the zero-cost state is the vacuum state. Mutual information can be written as a sum of contributions produced by individual measurement outcomes. The expansion depends on whether the measurement operator corresponding to a given outcome detects the zero-cost state or not. In the optical domain, these two classes of operators can be interpreted as photon counting and phase-sensitive detection such as homodyning. The developed formalism substantially helps in deriving bounds on the attainable mutual information in the low-cost limit, reducing the dimensionality of the relevant Hilbert space to zero- and one-photon Fock states and mapping the geometry of the ensemble onto a set of fixed state vectors that are independent of the cost parameter.

In the case of individual measurements on a coherent state ensemble with a vanishing mean photon number, we derived bounds on the attainable photon information efficiency that go beyond those obtained assuming conventional detection techniques. Furthermore, superadditivity of accessible information has been analysed for words composed of the BPSK alphabet of coherent states. The mathematical structure of the derived general bound has been used to identify collective measurements that offer performance beyond individual detection. In the case when single photon states characterizing the words in the low-cost limit are orthogonal one can calculate an analytical value for the asymptotic maximum PIE and the design of the collective measurement is apparent. The remaining cases in which one cannot assume orthogonality are much more involved and in general one need to resort to numerical optimization. We presented experimental schemes based on linear optics, homodyne detection, and photon counting which in the low-cost limit offer superadditive enhancement over the single-symbol measurements equal to 2.49% and 3.40% respectively for two- and three-symbol measurements. Their performance for finite mean photon number has been analyzed using as a benchmark the minimum-error measurement on single symbols, which is known to maximize the mutual information for individual detection. Interestingly, further optimization of collective detection requires a careful adjustment of the measurement to the value of cost parameter.

The developed mathematical formalism can find applications not only in standard classical communication but also in private communication, where the private communication rate can be expressed by the difference of the mutual informations between sender and receiver and the eavesdropper [57, 58]. An extension of our methods in line with [19] to tackle von Neumann or Renyi entropies may allow to analyze low-cost limits of various quantum information quantities like Holevo information [1], coherent information [59, 60] or other quantities, relevant in quantum key distribution.

Acknowledgments

We acknowledge insightful discussions with F E Becerra, R García-Patrón, S Guha and M G A Paris. This work is part of the project ‘Quantum Optical Communication Systems’ carried out within the TEAM programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund.

Appendix A. Photon information efficiency for displaced thermal states

The treatment of individual measurements performed on coherent states presented in section 4 can be generalized to the case of displaced thermal states

$$\hat{\rho}_j = \hat{D}(\gamma_j \zeta) \hat{\rho}_{\text{th}} \hat{D}^\dagger(\gamma_j \zeta), \quad (\text{A1})$$

where $\hat{D}(\gamma_j \zeta)$ is the displacement operator and

$$\hat{\rho}_{\text{th}} = \frac{1}{1 + n_b} \sum_{n=0}^{\infty} \left(\frac{n_b}{1 + n_b} \right)^n |n\rangle \langle n| \quad (\text{A2})$$

is a thermal state with the mean photon number n_b , which plays the role of the zero-cost state $\hat{\rho}^{(0)} = \hat{\rho}_{\text{th}}$. Such states are produced e.g. as a result of propagation of coherent states through a phase-invariant Gaussian channel with excess noise.

In the present case the zero-cost state is mixed and has support on the entire bosonic Hilbert space. Thus all POVM operators \hat{Q}_r need to be type- \mathcal{Z} . For a displaced thermal state, the SLD difference \hat{D}_j at $\zeta = 0$ defined in equation (16) reads

$$\hat{D}_j = \frac{2}{1 + 2n_b} [(\gamma_j - \gamma_{\text{ens}}) \hat{a}^\dagger + (\gamma_j^* - \gamma_{\text{ens}}^*) \hat{a}], \quad (\text{A3})$$

where \hat{a}^\dagger and \hat{a} denote respectively the creation and the annihilation operators of the bosonic mode.

Inserting equation (A3) into (19) yields an upper bound on mutual information I in the low cost limit in the form

$$\begin{aligned} I &= \sum_r I_r \leq \frac{\zeta^2}{2} \sum_j p_j \text{Tr}[\hat{D}_j^2 \hat{\rho}^{(0)}] \\ &= \frac{2\zeta^2}{(1 + 2n_b)^2} \sum_j p_j |\gamma_j - \gamma_{\text{ens}}|^2 \text{Tr}[\hat{a}^\dagger \hat{a} \varrho^{(0)} + \hat{a} \hat{a}^\dagger \varrho^{(0)}] \\ &= \frac{2\bar{n}}{1 + 2n_b}, \end{aligned} \quad (\text{A4})$$

where we have used $\text{Tr}[(\hat{a}^\dagger)^2 \rho^{(0)}] = 0$ and $\text{Tr}[\hat{a}^2 \rho^{(0)}] = 0$ as well as the definition of the mean photon number from equation (26). Note that equation (A4) is identical with the first order expansion of the Shannon–Hartley formula [11] for the scenario with excess noise

$$I_{\text{SH}} = \frac{1}{2} \log \left(1 + \frac{4\bar{n}}{1 + 2n_b} \right). \quad (\text{A5})$$

Appendix B. Homodyne measurement schemes

The statistics of outcomes for shot-noise limited homodyne detection of the real quadrature x on a coherent state with a complex amplitude ζ is given by a Gaussian probability distribution

$$p_{\text{hom}}(x|\zeta) = \frac{1}{\sqrt{\pi}} \exp[-(x - \sqrt{2} \text{Re } \zeta)^2]. \quad (\text{B1})$$

Mutual information for homodyne detection of the BPSK ensemble of coherent states with real amplitudes $|\zeta\rangle$ and $|\zeta\rangle$ can be obtained by plugging equation (B1) into (4) and (5)

$$I_{\text{BPSK/hom}} = 4\zeta^2 - \frac{e^{-2\zeta^2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx e^{-x^2} \cosh(2\sqrt{2}x\zeta) \log[\cosh(2\sqrt{2}x\zeta)], \quad (\text{B2})$$

with the summation over results exchanged to integration. Expanding all the functions in the second term up to the fourth order in ζ yields

$$I_{\text{BPSK/hom}} \approx 4\zeta^2 - \frac{1 - 2\zeta^2 + 2\zeta^4}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx e^{-x^2} \left(4x^2\zeta^2 + \frac{32}{3}x^4\zeta^4 \right) \approx 2\zeta^2 - 4\zeta^4. \quad (\text{B3})$$

The rightmost expression coincides with the power series expansion of the Shannon–Hartley limit given in equation (38).

The statistics outcomes on the single photon detector (SPD) deployed in the measurement scheme discussed in section 5.1 and depicted in figure 3 with an optional displacement β is given by

$$p_{\text{SPD}}(k|\zeta') = \begin{cases} e^{-|\zeta'+\beta|^2}, & k = 0 \\ 1 - e^{-|\zeta'+\beta|^2}, & k = 1, \end{cases} \quad (\text{B4})$$

where $k = 0, 1$ indicates respectively if any photons were registered or not and ζ' is the field amplitude before the displacement operation. The probabilities of possible outcomes in the two-symbol measurement design of figure 3 are then given by

$$p_{x,k|j} = p_{\text{hom}}\left(x|((-1)^{j_1} + (-1)^{j_2})\frac{\zeta}{\sqrt{2}}\right)p_{\text{SPD}}\left(k|((-1)^{j_1} - (-1)^{j_2})\frac{\zeta}{\sqrt{2}}\right). \quad (\text{B5})$$

The mutual information can be written as

$$I_2 = H(X, K) - \sum_j p_j \{H[p_{\text{SPD}}(0|j)] + h[p_{\text{hom}}(x|j)]\}, \quad (\text{B6})$$

where $H(X, K)$ denotes the joint entropy of the outcomes at the homodyne detector and the SPD, H is the entropy of a binary random variable defined in equation (37) and $h[p(x)] = -\int_{-\infty}^{\infty} dx p(x) \log p(x)$.

In the case of a pure SPD measurement without displacement ($\beta = 0$) one can calculate mutual information by expanding $\log p_{x,0}$ appearing in the expression for $H(X, K)$ up to the second order in ζ . Taking the limit of $\zeta^2 = \bar{n} \rightarrow 0$ in the resulting expression yields

$$I_2 \approx 4\bar{n}(1 - u) - 2\bar{n}u \log u = 2\bar{n}[2 + f(u)]. \quad (\text{B7})$$

The derived expression is equal to the mutual information of the two-symbol scenario of equation (52) and is maximized by $u = 1/e^3$. The advantage is 2.49% compared to the individual measurement.

ORCID iDs

Ludwig Kunz  <https://orcid.org/0000-0002-0613-7070>

Marcin Jarzyna  <https://orcid.org/0000-0001-9989-1648>

Wojciech Zwoleński  <https://orcid.org/0000-0001-7616-0404>

Konrad Banaszek  <https://orcid.org/0000-0002-5389-6897>

References

- [1] Holevo A S 1973 *Probl. Peredachi Inf.* **9** 3–11
- [2] Hausladen P, Jozsa R, Schumacher B, Westmoreland M and Wootters W K 1996 *Phys. Rev. A* **54** 1869–76
- [3] Holevo A S 1998 *IEEE Trans. Inf. Theory* **44** 269–73
- [4] Guha S and Wilde M M 2012 Polar coding to achieve the holevo capacity of a pure-loss optical channel *2012 IEEE Int. Symp. on Information Theory Proc.* pp 546–50
- [5] Wilde M M, Guha S, Tan S and Lloyd S 2012 Explicit capacity-achieving receivers for optical communication and quantum reading *2012 IEEE Int. Symp. on Information Theory Proc.* pp 551–5
- [6] Holevo A S 1979 *Probl. Peredachi Inf.* **15** 3–11
- [7] Peres A and Wootters W K 1991 *Phys. Rev. Lett.* **66** 1119–22
- [8] Sasaki M, Kato K, Izutsu M and Hirota O 1998 *Phys. Rev. A* **58** 146–58
- [9] Buck J R, van Enk S J and Fuchs C A 2000 *Phys. Rev. A* **61** 032309
- [10] Guha S 2011 *Phys. Rev. Lett.* **106** 240502
- [11] Banaszek K, Kunz L, Jachura M and Jarzyna M 2020 Quantum limits in optical communications arXiv:2002.05766
- [12] Banaszek K and Jachura M 2017 Structured optical receivers for efficient deep-space communication *Proc. 2017 IEEE Int. Conf. on Space Optical Systems and Applications (ICSOS) (Naha, Okinawa, Japan)* pp 34–7
- [13] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
- [14] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North Holland)
- [15] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439–43
- [16] Barndorff-Nielsen O E and Gill R D 2000 *J. Phys. A: Math. Gen.* **33** 4481–90
- [17] Liu J, Chen J, Jing X X and Wang X 2016 *J. Phys. A: Math. Theor.* **49** 275302
- [18] Šafránek D 2017 *Phys. Rev. A* **95** 052320
- [19] Czajkowski J, Jarzyna M and Demkowicz-Dobrzański R 2017 *New J. Phys.* **19** 073034
- [20] Caves C M and Drummond P D 1994 *Rev. Mod. Phys.* **66** 481–537
- [21] Shapiro J H 2009 *IEEE J. Sel. Top. Quant. Elec.* **15** 1547–69
- [22] Shannon C E 1949 *Proc. IRE* **37** 10–21
- [23] Verdu S 2002 *IEEE Trans. Inf. Theor.* **48** 1319–43
- [24] Guo D, Shamai S and Verdu S 2005 *IEEE Trans. Inf. Theor.* **51** 1261–82
- [25] Banaszek K, Kunz L, Jarzyna M and Jachura M 2019 Approaching the ultimate capacity limit in deep-space optical communication *Proc. SPIE* **10910** 109100A
- [26] Papen G C and Blahut R E 2019 *Lightwave Communications* (Cambridge: Cambridge University Press)
- [27] Davies E 1978 *IEEE Trans. Inf. Theor.* **24** 596–9
- [28] Sasaki M, Barnett S M, Jozsa R, Osaki M and Hirota O 1999 *Phys. Rev. A* **59** 3325–35
- [29] Waseda A, Sasaki M, Takeoka M, Fujiwara M, Toyoshima M and Assalini A 2011 *J. Opt. Commun. Netw.* **3** 514–21
- [30] Moision B and Farr W 2014 *IPN Prog. Rep.* **42**–199 1–10

- [31] Osaki M, Hirota O and Ban M 1998 *J. Mod. Opt.* **45** 269–82
- [32] Tomassoni N and Paris M G A 2008 *Phys. Lett. A* **373** 61–4
- [33] Dolinar S J 1973 *Q. Prog. Rep.* **111** 115
- [34] Cook R L, Martin P J and Geremia J M 2007 *Nature* **446** 774–7
- [35] Sych D and Leuchs G 2016 *Phys. Rev. Lett.* **117** 200501
- [36] Wittmann C, Takeoka M, Cassemiro K N, Sasaki M, Leuchs G and Andersen U L 2008 *Phys. Rev. Lett.* **101** 210501
- [37] Takeoka M and Sasaki M 2008 *Phys. Rev. A* **78** 022320
- [38] Wittmann C, Andersen U L, Takeoka M, Sych D and Leuchs G 2010 *Phys. Rev. Lett.* **104** 100505
- [39] Tsujino K, Fukuda D, Fujii G, Inoue S, Fujiwara M, Takeoka M and Sasaki M 2010 *Opt. Express* **18** 8107–14
- [40] Tsujino K, Fukuda D, Fujii G, Inoue S, Fujiwara M, Takeoka M and Sasaki M 2011 *Phys. Rev. Lett.* **106** 250503
- [41] DiMario M T and Becerra F E 2018 *Phys. Rev. Lett.* **121** 023603
- [42] DiMario M T, Kunz L, Banaszek K and Becerra F E 2019 *NPJ Quantum Inf.* **5** 65
- [43] Han R, Bergou J A and Leuchs G 2018 *New J. Phys.* **20** 043005
- [44] Becerra F E, Fan J, Baumgartner G, Polyakov S V, Goldhar J, Kosloski J T and Migdall A 2011 *Phys. Rev. A* **84** 062324
- [45] Izumi S, Takeoka M, Fujiwara M, Pozza N D, Assalini A, Ema K and Sasaki M 2012 *Phys. Rev. A* **86** 042328
- [46] Becerra F E, Fan J, Baumgartner G, Goldhar J, Kosloski J T and Migdall A 2013 *Nat. Photon.* **7** 147–52
- [47] Becerra F E, Fan J and Migdall A 2015 *Nat. Photon.* **9** 48–53
- [48] Ferdinand A R, DiMario M T and Becerra F E 2017 *NPJ Quantum Inf.* **3** 43
- [49] DiMario M T, Carrasco E, Jackson R A and Becerra F E 2018 *J. Opt. Soc. Am. B* **35** 568–74
- [50] Izumi S, Neergard-Nielsen J S, Miki S, Terai H and Andersen U L 2020 Experimental demonstration of a quantum receiver beating the standard quantum limit at the telecom wavelength arXiv:2001.05902
- [51] Chefles A 2000 *Contemp. Phys.* **41** 401–24
- [52] Kilin S Y 2001 Quanta and information *Progress in Optics* ed E Wolf vol 42 (Amsterdam: Elsevier) pp 1–91
- [53] Barnett S M and Croke S 2009 *Adv. Opt. Photonics* **1** 238–78
- [54] Horadam K J 2007 *Hadamard Matrices and their Applications* (Princeton, NJ: Princeton University Press)
- [55] Klimek A, Jachura M, Wasilewski W and Banaszek K 2016 *J. Mod. Opt.* **63** 2074–80
- [56] Jarzyna M, Kuszaj P and Banaszek K 2015 *Opt. Express* **23** 3170–5
- [57] Wilde M M 2013 *Quantum Information Theory* (Cambridge: Cambridge University Press)
- [58] Devetak I 2005 *IEEE Trans. Inf. Theor.* **55** 44–55
- [59] Schumacher B and Nielsen M A 1996 *Phys. Rev. A* **54** 2629–35
- [60] Lloyd S 1997 *Phys. Rev. A* **55** 1613–22