

## Purification of noisy quantum measurements

Michele Dall'Arno,<sup>1,2</sup> Giacomo Mauro D'Ariano,<sup>1,2</sup> and Massimiliano F. Sacchi<sup>1,3</sup>

<sup>1</sup>*Quit Group, Dipartimento di Fisica "A. Volta," via Bassi 6, I-27100 Pavia, Italy*

<sup>2</sup>*Istituto Nazionale di Fisica Nucleare, Gruppo IV, via Bassi 6, I-27100 Pavia, Italy*

<sup>3</sup>*Istituto di Fotonica e Nanotecnologie (IFN-CNR), Piazza Leonardo da Vinci 32, I-20133 Milano, Italy*

(Received 24 June 2010; published 18 October 2010)

We consider the problem of improving noisy quantum measurements by suitable preprocessing strategies making many noisy detectors equivalent to a single ideal detector. For observables pertaining to finite-dimensional systems (e.g., qubits or spins) we consider preprocessing strategies that are reminiscent of quantum error correction procedures and allow one to perfectly measure an observable on a single quantum system for increasing number of inefficient detectors. For measurements of observables with an unbounded spectrum (e.g., photon number and homodyne and heterodyne detection), the purification of noisy quantum measurements can be achieved by preamplification as suggested by Yuen [*Opt. Lett.* **12**, 789 (1987)].

DOI: [10.1103/PhysRevA.82.042315](https://doi.org/10.1103/PhysRevA.82.042315)

PACS number(s): 03.67.Pp, 03.65.Ta, 42.50.Dv

### I. INTRODUCTION

In many situations it is necessary to measure an observable in the presence of noise, e.g., when transmitting a quantum state through a noisy quantum channel that degrades it exponentially versus distance, corresponding to a degradation of the measurement.

A number of figures of merit can be used to characterize the noise of nonideal measurements. An example of such figures of merit is the variance of the outcomes distribution. An extensive analysis of the variance affecting quantum measurements has been done, for example, in Ref. [1]. In a communication scenario, a relevant figure of merit is represented by the mutual information between the measurement outcomes and the input alphabet encoded on an ensemble of states. The problem of how much classical information can be extracted from a quantum system was first deeply discussed by Holevo [2], who provided bounds on the accessible information, and then revisited in the framework of quantum information by Schumacher *et al.* [3]. A further figure of merit is the average probability of correctly distinguishing input states picked up from a given ensemble. This is one of the first problems faced by quantum estimation theory and has been addressed extensively in the literature [2,4–6]. Finally, another example of figure of merit is a suitable distance between the noisy and the ideal outcomes probability for fixed input states.

In this article, we consider the situation where  $N$  identical preparations of the state  $\rho_g$  belonging to some ensemble  $\mathcal{S} = \{(\rho_g, \rho_g)\}$  are given. We are allowed to use  $M$  nonideal detectors, with  $M \geq N$ . Each detector is described by a positive operator-valued measure (POVM), namely a set of positive operators  $\{P'_i\}$ , which provides a resolution of the identity, i.e.,  $\sum_i P'_i = I$ . Each POVM element  $P'_i$  is the noisy version of an ideal POVM element  $P_i$ . A generic quantum channel  $\mathcal{R}$  is allowed to act on the  $N$  identical copies of the state  $\rho_g$  before the  $M$  noisy POVMs  $\{P'_i\}$  are measured, and generic classical postprocessing can be done on the outcomes of such measurements. Such a scheme of “purification” of noisy measurements is depicted in Fig. 1. We address the problem of optimizing the quantum channel  $\mathcal{R}$  in order to reduce the effect of noise affecting the POVMs  $\{P'_i\}$ . We approach the problem through the minimization of the variance

of the maximum likelihood estimator for the parameter  $g$  and through the maximization of the mutual information between  $g$  and the measurement outcomes.

Notice the analogy between quantum error correction schemes [7], as depicted in Fig. 2, and the purification of measurements. For error correction, the message is first encoded by gate  $\mathcal{R}$  into one of the carefully chosen code words, which is then (possibly) corrupted by the noisy communication channel  $\mathcal{E}$ . Finally, in gate  $\mathcal{D}$  some set of commuting Hermitian operators are measured over the corruption, the syndrome is used to perform error correction, and, finally, the recovered code word is decoded into the original message. For purification of measurements, we are allowed to encode the  $N$  identical copies of input state  $\rho_g$  through the channel  $\mathcal{R}$ , in a way similar to quantum error correction. The aim of such encoding is very different, since after that we are forced to perform  $M$  measurements with the same noisy POVM  $\{P'_i\}$ , which provide us just classical outcomes to be classically postprocessed. The limitation of the measurement purification versus error correction is that the decoding  $\mathcal{D}$  is restricted to classical outcomes only. The problem we are considering is also similar to the problem solved by entanglement purification protocols [8], since we are generally trying to recast the use of a number of noisy measurements to an effective use of a smaller number with less noise. The article is organized as follows. In Sec. II we specify the general problem to a qubit with isotropic noise, and then we face the optimization considering different figures of merit, in Sec. III we show how to minimize the measurement noise, while in Sec. IV we maximize the mutual information between the parameter describing the state and the outcomes of the POVMs. In Secs. V and VI, we consider observables with unbounded spectrum, for which the concept of amplification applies, and we review the scheme of Yuen [9] for purifying photodetectors (Sec. V) and homodyne and heterodyne detectors (Sec. VI). Finally, Sec. VII is devoted to conclusions.

### II. PURIFICATION OF QUBIT MEASUREMENTS

Let us specify the general problem we are considering. We are provided with  $N$  identical copies of the input state  $\rho_g$  of

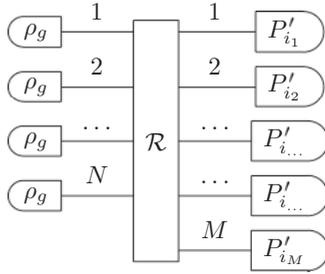


FIG. 1. Purification scheme for noisy quantum measurements.

dimension  $d$ . In what follows we will always suppose that the elements of the POVMs  $\{P_i\}$  and  $\{P'_i\}$  are  $d$ , which has been proved to be the optimal choice for  $d = 2$  [10], when the mutual information is optimized.<sup>1</sup> We suppose that each noisy element  $P'_i$  is obtained acting with the same channel  $\mathcal{E}$  on the corresponding element  $P_i$  of the ideal POVM

$$P'_i = \mathcal{E}^\vee(P_i), \quad (1)$$

where  $\mathcal{E}^\vee$  denotes the Heisenberg-picture version of the channel  $\mathcal{E}$ . Equation (1) shows that the ideal POVM  $\{P_i\}$  is “cleaner” than the noisy POVM  $\{P'_i\}$  in the sense of the partial ordering introduced in Ref. [12], as depicted in Fig. 3.

We consider a qubit (so  $d = 2$ ) parametrized as

$$\rho_{a,b} = \begin{pmatrix} a & b \\ b^* & 1-a \end{pmatrix}. \quad (2)$$

We are interested in the observable  $\sigma_z$ , and we suppose to have at our disposal  $M$  noisy POVMs  $\{P'_i\}$  of  $\sigma_z$ , i.e.,  $P_i = |i\rangle\langle i|$ . We assume a simple kind of noise acting on each POVM, i.e., the isotropic noise

$$\mathcal{E}^\vee(P_i) = \alpha P_i + \beta I, \quad (3)$$

so  $P'_i = \alpha|i\rangle\langle i| + \beta I$ .

We suppose we have an  $N = 1$  qubit state and consider as a purification channel  $\mathcal{R}$  the orthogonal cloning  $\mathcal{C}$ , with respect to the basis of eigenstates of the observable  $\sigma_z$ ,

$$\mathcal{C}(\rho) = \sum_{i=0,1} \langle i|\rho|i\rangle |i\rangle\langle i|^{\otimes M}. \quad (4)$$

The conditional probability  $p(\vec{i}|a,b)$  of obtaining outcomes  $\vec{i} = \{i_1, \dots, i_M\}$  given the state parametrized by  $a, b$  does not depend on  $b$  and can be explicitly written as

$$p(\vec{i}|a) = \text{Tr}[\mathcal{C}(\rho)\mathcal{E}^\vee(P_i)^{\otimes M}]. \quad (5)$$

<sup>1</sup>Indeed, for  $d > 2$  it has been shown in Ref. [11] that a measurement with number of outcomes larger than the dimension of the span of the input states can improve the mutual information.

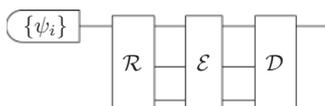


FIG. 2. Scheme for quantum error correction.

$$P'_i = \mathcal{E}(P_i)$$

FIG. 3. Noisy POVM element.

We substitute Eqs. (4) and (3) into Eq. (5) to obtain

$$p(\vec{i}|a) = \text{Tr}\{[a|0\rangle\langle 0|^{\otimes M} + (1-a)|1\rangle\langle 1|^{\otimes M}] \otimes_{j=1}^M (\alpha|i_j\rangle\langle i_j| + \beta I)\}. \quad (6)$$

We observe that the probability  $p(\vec{i}|a)$  depends only on the number of outcomes 0 and 1 in the measurement (i.e., not on their position). On defining such integers as  $M_0$  and  $M_1 = M - M_0$ , we obtain

$$p(M_1|a) = \binom{M}{M_1} [a(\alpha + \beta)^{M_0} \beta^{M_1} + (1-a)(\alpha + \beta)^{M_1} \beta^{M_0}]. \quad (7)$$

For the normalization condition of the POVM in Eq. (3) one has  $\alpha = 1 - 2\beta$ , so  $0 \leq \beta \leq \frac{1}{2}$ , and hence

$$p(M_1|a) = \binom{M}{M_1} \{a[(1-\beta)^{M_0} \beta^{M_1} - (1-\beta)^{M_1} \beta^{M_0}] + (1-\beta)^{M_1} \beta^{M_0}\}. \quad (8)$$

One can easily check the normalization of this probability, i.e.,  $\sum_{M_1=0}^M p(M_1|a) = 1$ . In the case of ideal measurements for which  $\beta = 0$ , the non-null probabilities are obtained just for  $M_1 = 0$  and for  $M_1 = M$ , namely

$$p(M_1 = 0|a) = a, \quad p(M_1 = M|a) = 1 - a, \quad (9)$$

whereas in the completely isotropic case (i.e.,  $\beta = \frac{1}{2}$ ) the probability  $p(M_1|a) = \binom{M}{M_1} (\frac{1}{2})^M$  is independent of  $a$ , namely no information can be obtained about the state. Note that also the coherent channel, widely used in encoding schemes for quantum error correction as [13]

$$\mathcal{C}'(\rho) = \sum_{i,j=0,1} \langle i|\rho|j\rangle |i\rangle\langle j|^{\otimes M}, \quad (10)$$

leads to the same probability distribution Eq. (8), since  $P'_i$  are diagonal on the  $\sigma_z$  basis.

### III. MINIMIZATION OF MEASUREMENT NOISE

We show how to apply the maximum likelihood (ML) criterion to obtain the optimal estimator for the expectation value of  $\sigma_z$ , by means of our measurement purification scheme. Our aim is to show an improvement of estimation in terms of variance by increasing the uses of the POVM.

The ML criterion provides the following estimator for  $a$  in the state Eq. (2):

$$a_{\text{ML}} = \arg \max_a \frac{1}{n} L(a|M_1), \quad (11)$$

where  $n$  is the number of (joint) outcomes (runs of the purification scheme depicted in Fig. 1),  $L(a|M_1)$  is the so-called log-likelihood functional

$$L(a|M_1) = \sum_{j=1}^n \log_2 p_j(M_1|a), \quad (12)$$

and  $p_j(M_1|a)$  denotes the conditional probability for the  $j$ th run. We observe that Eq. (11) is concave since the logarithm of a linear function is a concave function and the summation of concave functions is a concave function. To solve the ML problem, we employ the iterative numerical method described in Ref. [14]. First, we generate a large amount of data distributed according to Eq. (8) for some fixed value of  $a$  and  $\beta$ . Then, we fix some order zero approximation  $a_0$  for the estimator  $a_{ML}$ . Then, the first order correction is given by

$$a_1 = \frac{\left. \frac{\partial L(a|M_1)}{\partial a} \right|_{a=a_0}}{F(a_0)}, \quad (13)$$

where  $F(a)$  is the Fisher information

$$F(a) = \sum_{\vec{i}} \left[ \frac{\partial p(\vec{i}|a)}{\partial a} \right]^2 \frac{1}{p(\vec{i}|a)}, \quad (14)$$

which in our case is given by

$$F(a) = \sum_{M_1=0}^M \binom{M}{M_0}^2 \frac{[(1-\beta)^{M_0} \beta^{M_1} - (1-\beta)^{M_1} \beta^{M_0}]^2}{p(M_1|a)}. \quad (15)$$

The Fisher information measures the amount of information that the random variable  $M_1$  carries about the unknown parameter  $a$  on which the likelihood function depends. So the estimator to first order is  $a_0 + a_1$ , and the procedure can be iterated with this value as order zero approximation to obtain higher-order corrections. Obviously, the result is independent of the initial value  $a_0$ . For  $\beta$  not too big (say  $0 < \beta < \frac{1}{3}$ ), the algorithm converges in a few steps (say, less than 10). The variance on the ML estimator of a parameter satisfies the Cramer-Rao bound [15]

$$\sigma^2(a_{ML}) \geq \frac{1}{nF(a)}. \quad (16)$$

The bound in Eq. (16) is saturated if the number of data are large enough and the parameter is monodimensional (as in the present case). We numerically estimated the variance of the estimator  $a_{ML}$  in Eq. (11) by dividing the data into blocks, finding the estimator  $a_i$  for each block, and then calculating the variance of such estimators, namely

$$\sigma^2(a_{ML}) = \sum_i (a_i - a_{ML})^2. \quad (17)$$

In Fig. 4 we verified that the variance numerically saturates the bound in Eq. (16).

Figure 5 shows that the variance decreases as the number of POVMs used in parallel increases and upper and lower bounds for variance. To find the upper bound consider the function

$$f(M_1) = \frac{\frac{M_1}{M} + \beta - 1}{2\beta - 1}. \quad (18)$$

Note that  $f(M_1)$  is an unbiased estimator for the parameter  $a$ , since one has

$$\langle f(M_1) \rangle = \sum_{M_1=0}^M f(M_1) p(M_1|a) = a. \quad (19)$$

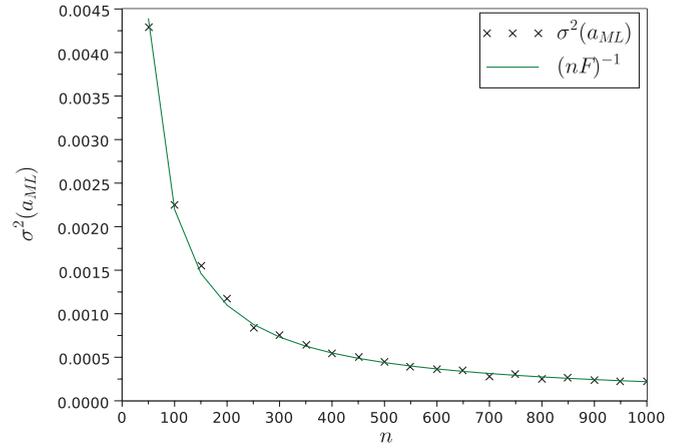


FIG. 4. (Color online) Variance  $\sigma^2(a_{ML})$  versus the number  $n$  of outcomes (runs) for parameter  $a = 0.75$ , measurement noise  $\beta = 0.25$ , and number of POVMs  $M = 10$ . The solid line represents the Cramer-Rao bound in Eq. (16).

The second moment is given by

$$\begin{aligned} \langle f(M_1)^2 \rangle &= \sum_{M_1=0}^M f(M_1)^2 p(M_1|a) \\ &= a + \frac{\beta(1-\beta)}{(1-2\beta)^2 M}. \end{aligned} \quad (20)$$

Thus, an upper bound for the variance on the parameter  $a$  is

$$\sigma^2(a_{ML}) \leq \left[ a - a^2 + \frac{\beta(1-\beta)}{(1-2\beta)^2 M} \right] \frac{1}{n}. \quad (21)$$

The lower bound for the variance is

$$\sigma^2(a_{ML}) \geq \frac{1}{nF(a)|_{\beta=0, M=1}} = \frac{a - a^2}{n}, \quad (22)$$

where the right-hand side of Eq. (22) corresponds to the use of the ideal detector on the original state.

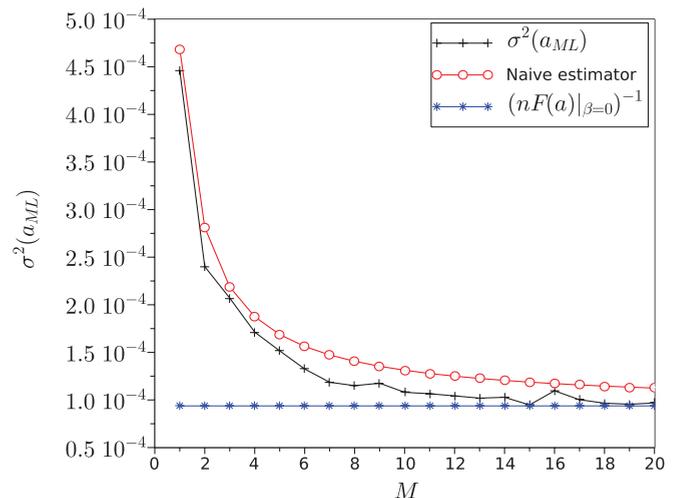


FIG. 5. (Color online) Variance  $\sigma^2(a_{ML})$  versus the number  $M$  of POVMs for parameter  $a = 0.75$ , measurement noise  $\beta = 0.25$ , and number of runs  $n = 2000$ . The upper and lower bounds correspond to Eqs. (21) and (22), respectively.

The computed variance shows a dependence on the parameter  $a$  similar that in Eq. (21), decreasing as  $1/M$  for  $a = 0$  or  $a = 1$ . The variance saturates the lower bound in Eq. (22), so the estimator of the parameter  $a$  is optimal.

#### IV. MAXIMIZATION OF MUTUAL INFORMATION

We consider now the mutual information as the figure of merit in the measurement purification scheme. We consider a qubit parametrized as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle. \quad (23)$$

The probability in Eq. (8) can be written as

$$p(M_1|\theta) = \binom{M}{M_1} \left[ [(1-\beta)^{M_0}\beta^{M_1} - (1-\beta)^{M_1}\beta^{M_0}] \cos^2\frac{\theta}{2} + (1-\beta)^{M_1}\beta^{M_0} \right], \quad (24)$$

independent of  $\phi$ .

In the following we suppose that the prior probability  $p(\theta, \phi)$  of having the input state in Eq. (23) is uniform, so the mutual information  $I(M_1 : \theta, \phi)$  between random variables  $\theta$  and  $\phi$  and random variable  $M_1$  is given by

$$I(M_1 : \theta) := \frac{1}{2} \int_0^\pi d\theta \sin\theta \sum_{M_1=0}^M p(M_1|\theta) \times \log_2 \left[ \frac{p(M_1|\theta)}{\frac{1}{2} \int_0^\pi d\theta \sin\theta p(M_1|\theta)} \right]. \quad (25)$$

The integral in the denominator gives

$$\int_0^\pi d\theta \sin\theta p(M_1|\theta) = \binom{M}{M_1} [(1-\beta)^{M_0}\beta^{M_1} + (1-\beta)^{M_1}\beta^{M_0}], \quad (26)$$

and a lengthy analytical form for Eq. (25) is provided in the Appendix. The mutual information  $I(M_1 : \theta)$  saturates the bound  $I(M_1 : \theta) \leq I(M_1 : \theta)|_{M=1, \beta=0} \simeq 0.279$  bit for increasing  $M$ . Note that the mutual information does not converge to 1 bit, since a continuous ‘‘alphabet’’ of states is allowed. The mutual information  $I(M_1 : \theta)$  saturates almost exponentially versus  $M$ , as shown by Fig. 6. This means that we are recasting the use of many noisy detectors to an effective use of a single ideal detector.

Let us consider in more detail the simplified case in which the only allowed input states are the up ( $\theta = \pi$ ) and down ( $\theta = 0$ ) eigenstates of  $\sigma_z$ . This simplification leads to two advantages: a much more tractable analytical form for the mutual information  $I_2(M_1 : \theta)$  and the possibility to make a comparison with classical postprocessing based on majority voting. The mutual information is given by

$$I_2(M_1 : \theta) = \sum_{M_1=0}^M \binom{M}{M_1} (1-\beta)^{M_1}\beta^{M_0} \times \log_2 \left[ \frac{2(1-\beta)^{M_1}\beta^{M_0}}{(1-\beta)^{M_0}\beta^{M_1} + (1-\beta)^{M_1}\beta^{M_0}} \right]. \quad (27)$$

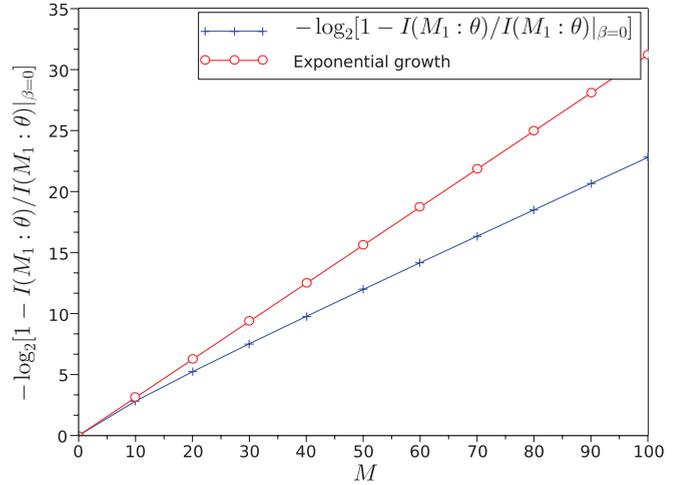


FIG. 6. (Color online) Function  $-\log_2[1 - \frac{I(M_1:\theta)}{I(M_1:\theta)|_{\beta=0}}]$  versus the number  $M$  of POVMs for measurement noise  $\beta = 0.25$ . The mutual information  $I(M_1 : \theta)$  is given by Eq. (A1).

Equation (27) behaves as expected for the ideal POVM case (i.e.,  $\beta = 0$ ), where  $I_2(M_1 : \theta) = 1$ , and for the completely isotropic POVM case (i.e.,  $\beta = \frac{1}{2}$ ), where  $I_2(M_1 : \theta) = 0$ . Finally, we investigate the optimal classical postprocessing to be applied on the  $M$  outcomes of the parallel noisy POVMs to maximize the mutual information. We simply argue that majority voting is close to the optimal postprocessing, as shown in Fig. 7. The gap between the binary mutual information  $I_2(M_1 : \theta)$  and that obtained with the majority-voting strategy could be explained by the fact that in general a number of POVM elements greater than the cardinality of the input alphabet can optimize the mutual information. In fact, Davies’s theorem [16] puts an upper bound of  $d^2$  on the number of POVM elements to optimize the mutual information for an alphabet of  $d$  linear independent pure states (see also Ref. [11]). The case of a two-states alphabet can be easily generalized to an alphabet of  $d$  orthogonal state  $\{|j\rangle\}$ , with  $j = 1, 2, \dots, d$ , and noisy POVM elements  $P_i = \alpha|i\rangle\langle i| + \frac{1-\alpha}{d}I$ . The conditional

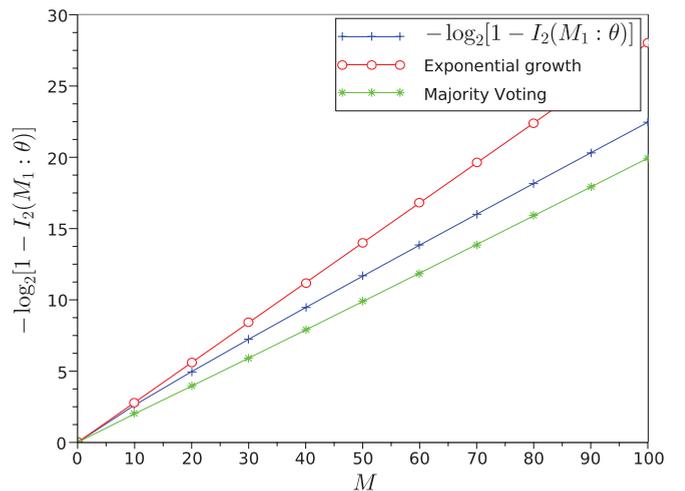


FIG. 7. (Color online) Function  $-\log_2[1 - I_2(M_1 : \theta)]$  versus the number  $M$  of POVMs for measurement noise  $\beta = 0.25$ .

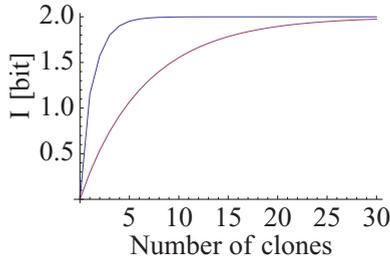


FIG. 8. (Color online) Mutual information for an alphabet of four equiprobable orthogonal states versus the number of purifying copies for measurement noise  $\alpha = 0.8$  (upper) and  $\alpha = 0.4$  (lower).

probability of the outcomes of  $M$  noisy measurements on  $M$  copies of  $|j\rangle$  is simply the multinomial

$$p(M_1, M_2, \dots, M_{d-1} | j) = \frac{M!}{M_1! M_2! \dots M_d} \frac{[(d-1)\alpha + 1]^{M_j} (1-\alpha)^{M-M_j}}{d^M}, \quad (28)$$

where  $M_l$  is the number of outcomes  $l$  in the string of  $M$  outcomes, and  $M_d = M - \sum_{j=1}^{d-1} M_j$ . The conditional probability allows one to evaluate the mutual information, and for increasing number of clones  $M$ , the noisy measurements are purified. In Fig. 8 we show the purification effect for an alphabet of four orthogonal equiprobable states and two different values of noise. As expected, for increasing value of  $M$  the mutual information approaches two bits.

## V. INEFFICIENT PHOTODETECTION

In the rest of the article we consider observables with an unbounded spectrum, for which the concept of amplification applies, and we review the scheme of Yuen [9] for improving noisy photodetectors and homodyne and heterodyne detectors. In the original proposal of Ref. [9] the signal-to-noise ratio improvement was studied for noisy measurements with preamplification assistance. By reviewing the results here, we explicitly consider the effect of amplification as a purification of the noisy POVMs and hence of the outcome probability distributions.

Light is revealed by exploiting its interaction with atoms, molecules, or electrons in a solid, and, essentially, each photon ionizes a single atom or promotes an electron to a conduction band, and the resulting charge is then amplified to produce a measurable pulse. In practice, however, available photodetectors do not ideally count all photons, and their performance is limited by a nonunit quantum efficiency  $\eta$ , namely only a fraction  $\eta$  of the incoming photons leads to an electric signal and ultimately to a *count*: some photons are either reflected from the surface of the detector or are absorbed without being transformed into electric pulses. Let us consider a light beam entering a photodetector of quantum efficiency  $\eta$ , i.e., a detector that transforms just a fraction  $\eta$  of the incoming light pulse into electric signal. We will focus our attention to the case of the radiation field excited in a stationary state of a single mode at frequency  $\omega$ . The Poissonian process of counting

then gives the following probability  $p_\eta(m)$  of revealing  $m$  photons [17]:

$$p_\eta(m) = \text{Tr} \left[ \rho : \frac{(\eta a^\dagger a)^m}{m!} \exp(-\eta a^\dagger a) : \right], \quad (29)$$

where  $\rho$  represents the quantum state of light, and  $: :$  denotes the normal ordering of field operators.

Using the identities

$$:(a^\dagger a)^n: = (a^\dagger)^n a^n = a^\dagger a (a^\dagger a - 1) \dots (a^\dagger a - n + 1), \quad (30)$$

$$:e^{-x a^\dagger a}: = \sum_{l=0}^{\infty} \frac{(-x)^l}{l!} (a^\dagger)^l a^l = (1-x)^{a^\dagger a}, \quad (31)$$

one obtains

$$p_\eta(m) = \sum_{n=m}^{\infty} \rho_{nn} \binom{n}{m} \eta^m (1-\eta)^{n-m}, \quad (32)$$

where

$$\rho_{nn} \equiv \langle n | \rho | n \rangle = p_{\eta=1}(n). \quad (33)$$

Hence, for unit quantum efficiency a photodetector measures the photon number distribution of the state, whereas for nonunit quantum efficiency the output distribution of counts is given by a Bernoulli convolution of the ideal distribution.

The outcome distribution in Eq. (32) can be equivalently described by means of a simple model in which the realistic photodetector is replaced with an ideal photodetector preceded by a beam splitter of transmissivity  $\tau \equiv \eta$ . The reflected mode is absorbed, whereas the transmitted mode is photodetected with unit quantum efficiency. In order to obtain the probability of measuring  $m$  clicks, note that, apart from trivial phase changes, a beam splitter of transmissivity  $\tau$  affects the unitary transformation of fields

$$\begin{pmatrix} c \\ d \end{pmatrix} \equiv U_\tau^\dagger \begin{pmatrix} a \\ b \end{pmatrix} U_\tau = \begin{pmatrix} \sqrt{\tau} & -\sqrt{1-\tau} \\ \sqrt{1-\tau} & \sqrt{\tau} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \quad (34)$$

where all field modes are considered at the same frequency. Hence, the output mode  $c$  hitting the detector is given by the linear combination

$$c = \sqrt{\tau} a - \sqrt{1-\tau} b, \quad (35)$$

and the probability of counts reads

$$p_\tau(m) = \text{Tr}[U_\tau(\rho \otimes |0\rangle\langle 0|)U_\tau^\dagger |m\rangle\langle m| \otimes 1] = \sum_{n=m}^{\infty} \rho_{nn} \binom{n}{m} (1-\tau)^{n-m} \tau^m. \quad (36)$$

Equation (32) is then reproduced for  $\tau = \eta$ . We conclude that a photodetector of quantum efficiency  $\eta$  is equivalent to a perfect photodetector preceded by a beam splitter of transmissivity  $\eta$  which accounts for the overall losses of the detection process. According to Eq. (32), the POVM describing the inefficient photodetector can be written as

$$\Pi_\eta(m) = \binom{a^\dagger a}{m} \eta^m (1-\eta)^{a^\dagger a - m}, \quad (37)$$

such that  $p_\eta(m) = \text{Tr}[\rho \Pi_\eta(m)]$ . The random variable  $m$ , suitably rescaled by  $\eta$ , provides an estimator of the average photon number  $\langle a^\dagger a \rangle = \text{Tr}[\rho a^\dagger a]$ , since one has

$$\sum_{m=0}^{\infty} \frac{m}{\eta} \Pi_\eta(m) = a^\dagger a. \quad (38)$$

In order to evaluate the second moment of the probability, one uses the identity

$$\sum_{m=0}^{\infty} \left(\frac{m}{\eta}\right)^2 \Pi_\eta(m) = (a^\dagger a)^2 + \frac{1-\eta}{\eta} a^\dagger a, \quad (39)$$

and hence the inefficient measurement is affected by the added noise  $\frac{1-\eta}{\eta} \langle a^\dagger a \rangle$ , with respect to the ideal intrinsic noise  $\Delta(a^\dagger a)^2 \equiv \langle (a^\dagger a)^2 \rangle - \langle a^\dagger a \rangle^2$ .

In the following we show that an ideal photon-number amplifier can arbitrarily reduce the added noise of the inefficient measurement for increasing gain. The ideal photon-number amplification map is given by [18–20]

$$a^\dagger a \longrightarrow \hat{V}^\dagger a^\dagger a \hat{V} = g a^\dagger a, \quad (40)$$

where  $g$  is an integer and  $\hat{V}$  is the isometry

$$\hat{V} = \sum_{n=0}^{\infty} |gn\rangle \langle n|. \quad (41)$$

The preamplified POVM is simply given by

$$\Pi_\eta^{(g)}(m) = \binom{ga^\dagger a}{m} \eta^m (1-\eta)^{ga^\dagger a - m}. \quad (42)$$

The estimator of the average photon number is now  $m/(g\eta)$ , and the second moment is given by

$$\sum_{m=0}^{\infty} \left(\frac{m}{g\eta}\right)^2 \Pi_\eta^{(g)}(m) = (a^\dagger a)^2 + \frac{1-\eta}{g\eta} a^\dagger a. \quad (43)$$

Clearly, for  $g \rightarrow \infty$ , the added noise is completely removed for any value of the quantum efficiency  $\eta$ .

We note that the ideal photon-number amplifier is so effective that, indeed, even a preamplified heterodyne detection provides the ideal photon number distribution for increasing gain, as shown in Ref. [21].

## VI. INEFFICIENT CONTINUOUS VARIABLE MEASUREMENTS

### A. Homodyne detection

The balanced homodyne detector provides the measurement of the quadrature of the field

$$X_\varphi = \frac{a^\dagger e^{i\varphi} + a e^{-i\varphi}}{2}. \quad (44)$$

It was proposed by Yuen and Chan [22] and subsequently experimentally demonstrated by Abbas, Chan, and Yee [23]. The signal mode  $a$  interferes with a strong laser beam mode  $b$  in a balanced 50:50 beam splitter. The mode  $b$  is the so-called *local oscillator* (LO) mode of the detector. It operates at the same frequency of  $a$  and is excited by the laser in a strong coherent state  $|z\rangle$ . Since in all experiments that use homodyne detectors the signal and the LO beams are generated

by a common source, we assume that they have a fixed phase relation. In this case the LO phase provides a reference for the quadrature measurement, namely we identify the phase of the LO with the phase difference between the two modes. By tuning  $\varphi = \arg z$  we can measure the quadrature  $X_\varphi$  at arbitrary phase.

Behind the beam splitter, the two modes are detected by two identical photodetectors (usually linear avalanche photodiodes), and finally the difference of photocurrents at zero frequency is electronically processed. In the strong-LO limit  $|z| \rightarrow \infty$ , the homodyne detector is described by the POVM

$$\Pi(x) = \int_{-\infty}^{+\infty} \frac{d\lambda}{2\pi} \exp[i\lambda(X_\varphi - x)] = |x\rangle_{\varphi\varphi} \langle x|, \quad (45)$$

namely the projector on the eigenstate of the quadrature  $X_\varphi$  with eigenvalue  $x$ . In conclusion, the balanced homodyne detector achieves the ideal measurement of the quadrature  $X_\varphi$  in the strong LO limit. In this limit, the probability distribution of the output photocurrent approaches exactly the probability distribution  $p(x, \varphi) = {}_\varphi \langle x | \rho | x \rangle_\varphi$  of the quadrature  $X_\varphi$  for any state  $\rho$  of the signal mode  $a$ .

It is easy to take into account nonunit quantum efficiency at detectors. The POVM is obtained by replacing

$$X_\varphi \rightarrow X_\varphi + \sqrt{\frac{1-\eta}{2\eta}} (u_\varphi + v_\varphi) \quad (46)$$

in Eq. (45) with  $w_\varphi = (w^\dagger e^{i\varphi} + w e^{-i\varphi})/2$ , where  $w = u, v$  denotes the vacuum modes of the two inefficient photodetectors. By tracing the vacuum modes  $u$  and  $v$ , one obtains

$$\begin{aligned} \Pi_\eta(x) &= \int_{-\infty}^{+\infty} \frac{d\lambda}{2\pi} e^{i\lambda(X_\varphi - x)} |\langle 0 | e^{i\lambda \sqrt{\frac{1-\eta}{2\eta}} u_\varphi} | 0 \rangle|^2 \\ &= \int_{-\infty}^{+\infty} \frac{d\lambda}{2\pi} e^{i\lambda(X_\varphi - x)} e^{-\lambda^2 \frac{1-\eta}{8\eta}} \\ &= \frac{1}{\sqrt{2\pi \Delta_\eta^2}} \exp\left[-\frac{(x - X_\varphi)^2}{2\Delta_\eta^2}\right] \\ &= \frac{1}{\sqrt{2\pi \Delta_\eta^2}} \int_{-\infty}^{+\infty} dx' e^{-\frac{1}{2\Delta_\eta^2}(x-x')^2} |x'\rangle_{\varphi\varphi} \langle x'|, \end{aligned} \quad (47)$$

where

$$\Delta_\eta^2 = \frac{1-\eta}{4\eta}. \quad (48)$$

Thus the noisy POVM, and in turn the probability distribution of the output photocurrent, are just the Gaussian convolution of the ideal ones with rms  $\Delta_\eta = \sqrt{(1-\eta)/(4\eta)}$ .

In the following we show that the added noise of the inefficient homodyne detector can be removed by amplifying the signal by means of a phase-sensitive amplifier. This amplifier is described by the squeezing operator

$$S(\xi) = \exp\left[\frac{1}{2}(\xi a^{\dagger 2} - \xi^* a^2)\right] \quad (49)$$

and performs the mode transformation

$$S^\dagger(\xi) a S(\xi) = (\cosh |\xi|) a + \frac{\xi}{|\xi|} (\sinh |\xi|) a^\dagger. \quad (50)$$

For  $\xi = r e^{2i\varphi}$ , with  $r > 0$ , one has

$$S^\dagger(\xi)X_\varphi S(\xi) = e^r X_\varphi. \quad (51)$$

Hence, the effective POVM obtained by preprocessing  $\Pi_\eta(x)$  in Eq. (47) with the phase-sensitive amplification of  $X_\varphi$  is given by

$$\begin{aligned} \Pi_\eta^{(r)}(x) &= S^\dagger(r e^{2i\varphi})\Pi_\eta(x)S(r e^{2i\varphi}) \\ &= \frac{1}{\sqrt{2\pi\Delta_\eta^2}} \exp\left[-\frac{(x - e^r X_\varphi)^2}{2\Delta_\eta^2}\right]. \end{aligned} \quad (52)$$

Now, in order to obtain an unbiased measurement of  $X_\varphi$ , it is enough to rescale the outcome by  $e^r$ . On the other hand, the added noise with respect to the ideal measurement  $X_\varphi$  becomes equal to  $e^{-2r}\Delta_\eta^2$ , which can be made arbitrary small for increasing value of the squeezing parameter  $r$ .

### B. Heterodyne detection

Heterodyne detection allows to perform the joint measurement of two conjugated quadratures of the field [24,25].

A strong local oscillator at frequency  $\omega$  in a coherent state  $|\alpha\rangle$  hits a beam splitter with transmissivity  $\tau \rightarrow 1$  and with the coherent amplitude  $\alpha$  such that  $\gamma \equiv |\alpha|\sqrt{\tau(1-\tau)}$  is kept constant. If the output photocurrent is sampled at the intermediate frequency  $\omega_{\text{IF}}$ , just the field modes  $a$  and  $b$  at frequency  $\omega \pm \omega_{\text{IF}}$  are selected by the detector. Modes  $a$  and  $b$  are usually referred to as signal band and image band modes, respectively. In the strong LO limit, on tracing the LO mode, the output photocurrent  $I(\omega_{\text{IF}})$  rescaled by  $\gamma$  is equivalent to the complex operator

$$Z = \frac{I(\omega_{\text{IF}})}{\gamma} = a - b^\dagger, \quad (53)$$

where the arbitrary phases of modes have been suitably chosen. The heterodyne photocurrent  $Z$  is a normal operator, equivalent to a couple of commuting self-adjoint operators

$$Z = \text{Re}Z + i\text{Im}Z, \quad [Z, Z^\dagger] = [\text{Re}Z, \text{Im}Z] = 0. \quad (54)$$

The POVM of the detector is then given by the orthogonal (in Dirac sense) eigenvectors of  $Z$ .

In conventional heterodyne detection the image band mode is in the vacuum state and one is just interested in measuring the field mode  $a$ . In this case the POVM  $\Pi(z)$  is obtained on tracing on mode  $b$ , and one has the customary projectors on coherent states

$$\Pi(z) = \frac{1}{\pi} |z\rangle\langle z|, \quad (55)$$

with  $z \in \mathbb{C}$ . The coherent-state POVM provides the optimal joint measurement of conjugated quadratures of the field [4]. For a state  $\rho$ , the expectation value of any quadrature  $X_\varphi$  is obtained as

$$\langle X_\varphi \rangle = \text{Tr}[\rho X_\varphi] = \int_{\mathbb{C}} \frac{d^2\alpha}{\pi} \text{Re}(\alpha e^{-i\varphi}) \langle \alpha | \rho | \alpha \rangle. \quad (56)$$

The price to pay for jointly measuring noncommuting observables is an additional noise. The rms fluctuation is evaluated as follows:

$$\int_{\mathbb{C}} \frac{d^2\alpha}{\pi} [\text{Re}(\alpha e^{-i\varphi})]^2 \langle \alpha | \rho | \alpha \rangle - \langle X_\varphi \rangle^2 = \langle \Delta X_\varphi^2 \rangle + \frac{1}{4}, \quad (57)$$

where  $\langle \Delta X_\varphi^2 \rangle$  is the intrinsic noise, and the additional term is usually referred to as “the additional 3dB noise due to the joint measure” [26–28].

The effect of nonunit quantum efficiency can be taken into account in analogous way as in Sec. VIA for homodyne detection. The coherent-state POVM is replaced with the convolution

$$\Pi_\eta(z) = \int_{\mathbb{C}} \frac{d^2z'}{\pi \Delta_\eta^2} e^{-\frac{|z'-z|^2}{\Delta_\eta^2}} \frac{|z'\rangle\langle z'|}{\pi}, \quad (58)$$

where  $\Delta_\eta^2 = (1 - \eta)/\eta$ .

In the following we show that inefficient heterodyne detection can be purified by phase-insensitive amplification. Phase-insensitive amplification with (power) gain  $G$  amplifies the coherent amplitude of coherent states by  $\sqrt{G}$  at the expense of addition thermal photons  $\bar{n} = G - 1$ . Differing from phase-insensitive amplification, the physical process is not unitary but described by a completely positive map  $\mathcal{E}_G$ . Here, we just need the Heisenberg evolution of the projector on coherent states, which is simply given by the rescaling [29]

$$\mathcal{E}_G^\times(|\alpha\rangle\langle\alpha|) = \frac{1}{G} |G^{-1/2}\alpha\rangle\langle G^{-1/2}\alpha|. \quad (59)$$

It follows that under phase-insensitive preamplification the noisy heterodyne POVM (58) is replaced with

$$\Pi_\eta^{(G)}(z) = \int_{\mathbb{C}} \frac{d^2z'}{\pi \Delta_\eta^2} e^{-\frac{|\sqrt{G}z'-z|^2}{\Delta_\eta^2}} \frac{|z'\rangle\langle z'|}{\pi}. \quad (60)$$

Upon rescaling  $z \rightarrow G^{-1/2}z$ , one obtains

$$\begin{aligned} \Pi_\eta^{(G)}(z) &= \int_{\mathbb{C}} \frac{G d^2z'}{\pi \Delta_\eta^2} e^{-\frac{G|z'-z|^2}{\Delta_\eta^2}} \frac{|z'\rangle\langle z'|}{\pi} \\ &=_{G \rightarrow \infty} \frac{|z\rangle\langle z|}{\pi}, \end{aligned} \quad (61)$$

namely the noise due to quantum efficiency can be arbitrarily reduced for increasing value of the gain  $G$ . The effectiveness of preamplification in purifying heterodyne detection is more unexpected than the case of homodyne detection, since phase-insensitive amplification is not a unitary process.

### VII. CONCLUSION

In this article, we addressed the problem of optimizing a quantum channel acting before many parallel uses of a noisy POVM in order to purify the measurements, namely to achieve an effective measurement that is less noisy than the original ones. We first considered the purification of  $\sigma_z$  noisy measurements on qubits, by choosing the orthogonal cloning channel as a purification map. We found the maximum-likelihood estimator for  $\sigma_z$ , whose variance shows an almost exponential decay versus the number of POVMs. We also worked out

an analytic form for the mutual information between the state parameter and the outcomes of the POVMs, and here also an almost exponential improvement versus the number of POVMs has been found. We proved that naive majority voting is not the optimal classical postprocessing, since the maximum-likelihood approach gives a better estimator. For photodetection and continuous variable measurements as homodyne and heterodyne detection, the measurement purification can be achieved by preamplification, as early pointed out by Yuen [9].

We think that the relevant problem of purifying noisy quantum measurements will have a significant impact on the quantum information technology, in this same way as the decoherence problem.

#### ACKNOWLEDGMENTS

This work is supported by Italian Ministry of Education through PRIN 2008 and the European Community through the COQUIT project.

#### APPENDIX: DERIVATION OF THE MUTUAL INFORMATION IN EQ. (25)

We provide here an analytic form for Eq. (25), obtained substituting Eq. (8) for  $p(M_1|\theta)$ . One obtains

$$I(M_1 : \theta) = \frac{1}{32} \sum_{M_1=0}^M \binom{M}{M_1} \frac{c_1 A + c_2 B}{C}, \quad (\text{A1})$$

where

$$c_1 = (1 - \beta)^{2M} \beta^{4M_1}, \quad c_2 = (1 - \beta)^{4M_1} \beta^{2M}, \quad (\text{A2})$$

and

$$A = -16 \log_2[(1 - \beta)^{2M_1} \beta^M + (1 - \beta)^M \beta^{2M_1}] + 8 \log_2[(1 - \beta)^{2M} \beta^{4M_1}] + 16 - 8 \frac{1}{\ln(2)}, \quad (\text{A3})$$

$$B = +16 \log_2[(1 - \beta)^{2M_1} \beta^M + (1 - \beta)^M \beta^{2M_1}] - 8 \log_2[(1 - \beta)^{4M_1} \beta^{2M}] - 16 + 8 \frac{1}{\ln(2)}, \quad (\text{A4})$$

$$C = (1 - \beta)^{M+M_1} \beta^{3M_1} - (1 - \beta)^{3M_1} \beta^{M+M_1}. \quad (\text{A5})$$

- 
- [1] M. Ozawa, *Ann. Phys.* **311**, 350 (2004).  
[2] A. S. Holevo, *Problems Inform. Transmission* **9**, 110 (1973).  
[3] B. Schumacher, M. Westmoreland, and W. K. Wootters, *Phys. Rev. Lett.* **76**, 3452 (1996).  
[4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).  
[5] C. A. Fuchs, e-print [arXiv:quant-ph/9601020v1](https://arxiv.org/abs/quant-ph/9601020v1).  
[6] F. Buscemi and M. F. Sacchi, *Phys. Rev. A* **74**, 052320 (2006).  
[7] D. Gottesman, e-print [arXiv:quant-ph/0904.2557v1](https://arxiv.org/abs/quant-ph/0904.2557v1).  
[8] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, *Phys. Rev. Lett.* **76**, 722 (1996).  
[9] H. P. Yuen, *Opt. Lett.* **12**, 789 (1987).  
[10] L. B. Levitin, in *Quantum Communications and Measurement*, edited by V. P. Belavkin, O. Hirota, and R. L. Hudson (Plenum Press, New York, 1995), p. 439; A. Keil, e-print [arXiv:0809.0232](https://arxiv.org/abs/0809.0232) [quant-ph].  
[11] P. W. Shor, in *Quantum Communication, Computing, and Measurement 3*, edited by P. Tombesi and O. Hirota (Kluwer, New York, 2001), p. 107.  
[12] F. Buscemi, G. M. D'Ariano, M. Keyl, P. Perinotti, and R. Werner, *J. Math. Phys.* **46**, 082109 (2005).  
[13] P. W. Shor, *Phys. Rev. A* **52**, R2493 (1995).  
[14] C. R. Rao, *Linear Statistical Inference and Its Applications* (John Wiley & Sons, New York, 1976).  
[15] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).  
[16] E. B. Davies, *IEEE Inf. Theory* **24**, 596 (1978).  
[17] P. L. Kelley and W. H. Kleiner, *Phys. Rev.* **136**, A316 (1964).  
[18] H. P. Yuen, *Phys. Lett. A* **113**, 405 (1986).  
[19] G. M. D'Ariano, *Phys. Rev. A* **45**, 3224 (1992).  
[20] G. M. D'Ariano, C. Macchiavello, N. Sterpi, and H. P. Yuen, *Phys. Rev. A* **54**, 4712 (1996).  
[21] G. M. D'Ariano, M. F. Sacchi, and H. P. Yuen, *Int. J. Mod. Phys. B* **13**, 3069 (1999).  
[22] H. P. Yuen and V. W. S. Chan, *Opt. Lett.* **8**, 177 (1983).  
[23] G. L. Abbas, V. W. S. Chan, and S. T. Yee, *Opt. Lett.* **8**, 419 (1983); *IEEE J. Light. Tech.* **3**, 1110 (1985).  
[24] H. P. Yuen and J. H. Shapiro, *IEEE Trans. Inf. Theory* **24**, 657 (1978); **25**, 179 (1979).  
[25] H. P. Yuen and J. H. Shapiro, *IEEE Trans. Inf. Theory* **26**, 78 (1980).  
[26] E. Arthurs and J. L. Kelly, Jr., *Bell Syst. Tech. J.* **44**, 725 (1965).  
[27] H. P. Yuen, *Phys. Lett. A* **91**, 101 (1982).  
[28] E. Arthurs and M. S. Goodman, *Phys. Rev. Lett.* **60**, 2447 (1988).  
[29] C. M. Caves and P. D. Drummond, *Rev. Mod. Phys.* **66**, 481 (1994).