

## Quantum phase estimation with lossy interferometers

R. Demkowicz-Dobrzanski,<sup>1</sup> U. Dorner,<sup>2</sup> B. J. Smith,<sup>2,3</sup> J. S. Lundeen,<sup>2</sup> W. Wasilewski,<sup>4</sup> K. Banaszek,<sup>1</sup> and I. A. Walmsley<sup>2</sup>

<sup>1</sup>*Institute of Physics, Nicolaus Copernicus University, Grudziadzka 5, PL-87-100 Toruń, Poland*

<sup>2</sup>*Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*

<sup>3</sup>*Centre for Quantum Technologies, National University of Singapore, 117543 Singapore, Singapore*

<sup>4</sup>*Institute of Experimental Physics, University of Warsaw, Hoża 69, PL-00-681 Warsaw, Poland*

(Received 16 April 2009; published 24 July 2009)

We give a detailed discussion of optimal quantum states for optical two-mode interferometry in the presence of photon losses. We derive analytical formulae for the precision of phase estimation obtainable using quantum states of light with a definite photon number and prove that maximization of the precision is a convex optimization problem. The corresponding optimal precision, i.e., the lowest possible uncertainty, is shown to beat the standard quantum limit thus outperforming classical interferometry. Furthermore, we discuss more general inputs: states with indefinite photon number and states with photons distributed between distinguishable time bins. We prove that neither of these is helpful in improving phase estimation precision.

DOI: [10.1103/PhysRevA.80.013825](https://doi.org/10.1103/PhysRevA.80.013825)

PACS number(s): 42.50.St, 03.65.Ta, 06.20.Dk, 42.50.Lc

### I. INTRODUCTION

The strong sensitivity of certain quantum states to small variations of external parameters opens up great opportunities for devising high-precision measurements, e.g., of length and time, with unprecedented accuracy. A particularly important physical measurement technique is interferometry. Its numerous variations include Ramsey spectroscopy in atomic physics, optical interferometry in gravitational wave detectors, laser gyroscopes, and optical imaging to name but a few. Understanding limits on its performance in realistic situations under given resources is therefore of fundamental importance to metrology. In this paper we examine the fundamental limits of the precision of optical interferometry in the presence of photon losses for quantum states of light with definite photon number.

Optical interferometry aims to estimate the relative phase of two modes, or two “arms,” of the interferometer. This estimation process requires a certain amount of resources which is typically identified to be the number of photons,  $N$ , used for the measurement. The best precision which can be obtained using classical states of light scales like  $1/\sqrt{N}$ , the so-called standard quantum limit (SQL). Using nonclassical states of light this precision can be greatly improved, ideally leading to Heisenberg-limited scaling,  $1/N$  [1,2]. Indeed, recent years have seen many experimental proof-of-principle demonstrations of beating the SQL using quantum strategies in various interferometric setups [3–8]. Unfortunately, highly nonclassical states of light which potentially lead to Heisenberg-limited sensitivity are very fragile with respect to unwanted but unavoidable noise in experiments. In quantum-enhanced optical interferometry, the loss of photons is the most common and potentially the most devastating type of noise that one encounters. In particular it was noted that highly entangled quantum states, optimal for interferometry in the lossless case—the path-entangled states  $(|N,0\rangle + |0,N\rangle)/\sqrt{2}$ , or  $N00N$  states [9]—are extremely fragile. Even for moderate losses they are outperformed by purely classical states [10–16]. A different approach has been taken in [17], where the noise arising from imperfect preparation of a state has been investigated.

In [14], the first systematic approach was taken in order to determine the structure of optical states optimal for interferometry in the presence of losses. The best possible precision using input states with definite photon number was given. In this paper we elaborate and extend the ideas presented in [14]. Our treatment is based on general quantum measurement theory [18–21]. The quantity of interest is the lowest possible uncertainty attainable in parameter estimation, inversely proportional to the square root of the quantum Fisher information [20]. The quantum Fisher information depends only on the state of the system and not on the measurement procedure. We show that the optimization of the quantum Fisher information can be done effectively over the class of input states which have a definite photon number. Since these input states are subject to unavoidable photon losses they will degrade into mixed states and their suitability for phase estimation is compromised. Our optimization takes this into account yielding the most suitable input states in the presence of photon losses leading to the highest possible quantum Fisher information, and hence to the best possible precision. We give a detailed description of the noise model and calculate an analytic expression for the quantum Fisher information. The latter is shown to be a concave function on a convex set and therefore suited for efficient convex optimization methods. We numerically determine the optimal input states, compare them to alternative quantum and classical strategies, and show that they can beat the SQL. We note that the corresponding precision, which lies between the SQL and the Heisenberg limit (depending on the loss rates), defines the best possible precision for optical two-mode interferometry. In addition to this we discuss a measurement procedure that allows one to achieve the optimal precision, in terms of a positive operator-valued measure (POVM), and the possibility of using states with indefinite photon number or distinguishable photons. We show that neither of these generalizations improves the estimation precision, and consequently the state with definite photon number and indistinguishable photons are optimal.

The paper is organized as follows. In Sec. II a general scheme of quantum phase estimation is presented, and the

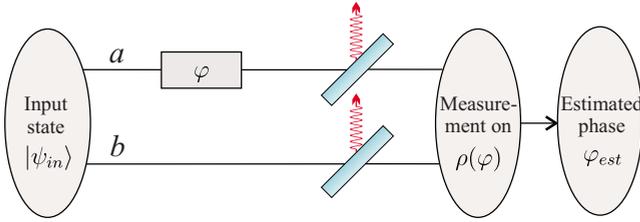


FIG. 1. (Color online) Interferometric phase estimation scheme. An input state is fed into an interferometer consisting of two channels. Channel  $a$  acquires a phase  $\varphi$  relative to channel  $b$ . Measurements are performed on the output state yielding an estimated value,  $\varphi_{est}$ , of the phase  $\varphi$ . The beam splitters symbolize photon losses.

notion of optimality is defined. In Sec. III we introduce the quantum Fisher information and discuss its most important properties. In Sec. IV we derive an explicit formula for the quantum Fisher information and prove that it is a concave function of input state parameters. In Sec. V we discuss the structure of the optimal states for interferometry and compare them to alternative strategies and states. In Sec. VI we discuss a measurement with which it is possible to achieve optimal precision. Finally, in Sec. VII we discuss possible generalizations of the considered quantum states, particularly states with indefinite photon number and the case when photons are distinguishable.

## II. PHASE ESTIMATION

We consider a general interferometer with two arms as shown in Fig. 1. A pure input state  $|\psi_{in}\rangle$  is fed into the interferometer and acquires a phase  $\varphi$  in the channel  $a$  relative to the channel  $b$ . Both channels, or “arms,” of the interferometer are subject to photon losses which can be modeled by fictitious beam splitters inserted at arbitrary locations in both channels. The output of the interferometer therefore needs to be described in general by a mixed state  $\rho(\varphi)$ . A measurement, represented by a positive operator-valued measure  $\{\Pi_i\}$  which defines a probability distribution for the measurement outcomes,

$$p(i|\varphi) = \text{Tr}[\Pi_i \rho(\varphi)], \quad (1)$$

is subsequently performed on the output state  $\rho(\varphi)$ . An estimated value of the true phase  $\varphi$  is obtained by applying an estimator that assigns to a particular measurement result  $i$  an estimated value  $\varphi_{est}(i)$ . We aim to estimate the phase  $\varphi$  as precisely as possible.

Two more elements need to be specified in order to make the problem of finding the optimal phase estimation strategy well defined: an *a priori* knowledge on the phase distribution  $p(\varphi)$  and a cost function  $C(\varphi, \varphi_{est})$  which can be seen as a measure for the uncertainty in the estimated phase. For phase estimation, the optimal choice of the input state, measurement and estimator is the one that minimizes the average cost function

$$\langle C \rangle = \int d\varphi p(\varphi) \sum_i p(i|\varphi) C(\varphi, \varphi_{est}(i)). \quad (2)$$

At this point two different approaches are most often pursued. In the *global approach* one assumes initial ignorance about the actual value of  $\varphi$ , which corresponds to the choice  $p(\varphi) = 1/2\pi$ . The solution of the problem then yields an estimation strategy which performs equally well irrespectively of the actual value of the estimated phase. In the *local approach*, on the other hand, the assumption is that the value of the actual phase lies in the vicinity of a known phase  $\varphi_0$ . More precisely, the *a priori* probability is chosen to be  $p(\varphi) = \delta(\varphi - \varphi_0)$ , while the estimator is required to be *locally unbiased* [22]

$$\frac{d}{d\varphi} \sum_i p(i|\varphi) \varphi_{est}(i) \Big|_{\varphi=\varphi_0} = 1. \quad (3)$$

The above condition is equivalent to a statement that the estimator will on average yield the true value of  $\varphi$  up to the first order in  $(\varphi - \varphi_0)$ . Notice that without local unbiasedness the estimation problem would be trivial (and also useless) since in order to minimize  $\langle C \rangle$  in Eq. (2), with  $p(\varphi) = \delta(\varphi - \varphi_0)$ , the optimal choice for the estimator would be simply  $\varphi_{est}(i) = \varphi_0$ , and the choice of the measurement would be irrelevant.

The local approach is useful when we are interested in small deviations of the phase from a known one. A significant advantage of the local approach over the global one is that for a natural choice of a quadratic cost function, there exist explicit lower bounds on  $\langle C \rangle$  based on the Fisher information (see Sec. III). In many practical situations these bounds are tight and the optimization over the measurement and the estimator can be avoided. Moreover, the local approach may also be useful in situations when there is no *a priori* knowledge on the phase  $\varphi$ . If many copies of a state are given, one can first perform a rough measurement (even not optimal) on a small fraction of copies in order to narrow down the range of potential values of  $\varphi$  so that they lie in a vicinity of a known phase  $\varphi_0$ , and then perform the optimal estimation using the local approach. This strategy will yield a high accuracy estimation for the global approach, without the need of optimizing the measurement and the estimator, since the rough measurement performed on small fraction of copies, even if not optimal, will not significantly influence the final accuracy [23]. See also [24] for discussion of interpolating regimes between the global and the local approaches.

## III. FISHER INFORMATION

In what follows we take the local approach. Since in this case we deal with small deviations of estimated phases from the true one, it is natural to choose  $C(\varphi, \varphi') = (\varphi - \varphi')^2$  as the cost function. Our goal is to find the optimal state, measurement and locally unbiased estimator minimizing the expression

$$(\delta\varphi)^2 = \sum_i p(i|\varphi) [\varphi_{est}(i) - \varphi]^2. \quad (4)$$

To minimize the standard deviation  $\delta\varphi$  given by the above formula we use an upper bound on  $\delta\varphi$  based on the Fisher information [18,25]. For a given POVM  $\{\Pi_i\}$  and state  $\rho(\varphi)$

defining the probabilities  $p(i|\varphi)$ , the Cramér-Rao inequality bounds the variance that can be obtained using any locally unbiased estimator,

$$(\delta\varphi)^2 \geq \frac{1}{F}, \quad (5)$$

where the Fisher information  $F$  is given by

$$F = \sum_i \frac{1}{p(i|\varphi)} \left( \frac{\partial p(i|\varphi)}{\partial \varphi} \right)^2. \quad (6)$$

If an experiment is repeated  $\nu$  times the bound reads

$$(\delta\varphi)^2 \geq \frac{1}{\nu F}. \quad (7)$$

For large  $\nu$  the Cramér-Rao bound is asymptotically achieved by the maximum likelihood estimator [18,20,25].

Optimization over the measurements yields the quantum Cramér-Rao bound [18–21]

$$(\delta\varphi)^2 \geq \frac{1}{F} \geq \frac{1}{F_Q}, \quad (8)$$

where the quantum Fisher information  $F_Q$  is given by

$$F_Q = \text{Tr}[\rho(\varphi)A^2]. \quad (9)$$

The Hermitian operator  $A$  is called the ‘‘symmetric logarithmic derivative’’ (SLD) and is implicitly defined via the relation

$$\frac{\partial \rho(\varphi)}{\partial \varphi} = \frac{1}{2}[A\rho(\varphi) + \rho(\varphi)A]. \quad (10)$$

In the eigenbasis of  $\rho(\varphi)$ ,  $A$  is given by

$$(A)_{ij} = \frac{2}{p_i + p_j} [\rho'(\varphi)]_{ij}, \quad (11)$$

where  $\rho'(\varphi) = \frac{\partial \rho(\varphi)}{\partial \varphi}$  and the  $p_i$  are the eigenvalues of  $\rho(\varphi)$  [whenever  $p_i + p_j = 0$  we set  $(A)_{ij} = 0$ ]. It has been shown that a measurement saturating the quantum Cramér-Rao bound exists and is given by a projective measurement on the eigenbasis of  $A$  [20,21].

For the sake of completeness we state some important properties of  $F_Q$  (see e.g., [18,26]):

(i) Let  $\rho(\varphi)$ ,  $\sigma(\varphi)$  be two density matrices supported on orthogonal subspaces,  $\mathcal{S}_{\rho(\varphi)} \perp \mathcal{S}_{\sigma(\varphi)}$ , which do not cease to be orthogonal for an infinitesimal change of  $\varphi$ , i.e.,  $(\mathcal{S}_{\rho(\varphi)} \cup \mathcal{S}_{\rho'(\varphi)}) \perp (\mathcal{S}_{\sigma(\varphi)} \cup \mathcal{S}_{\sigma'(\varphi)})$ , then  $F_Q$  is linear on the direct sum

$$F_Q[p\rho(\varphi) \oplus (1-p)\sigma(\varphi)] = pF_Q[\rho(\varphi)] + (1-p)F_Q[\sigma(\varphi)]. \quad (12)$$

(ii)  $F_Q$  is convex

$$F_Q[p\rho(\varphi) + (1-p)\sigma(\varphi)] \leq pF_Q[\rho(\varphi)] + (1-p)F_Q[\sigma(\varphi)]. \quad (13)$$

(iii) For pure states  $\rho(\varphi) = |\psi(\varphi)\rangle\langle\psi(\varphi)|$ ,  $F_Q$  reads

$$F_Q = 4[\langle\psi'(\varphi)|\psi'(\varphi)\rangle - |\langle\psi'(\varphi)|\psi(\varphi)\rangle|^2], \quad (14)$$

where  $|\psi'(\varphi)\rangle = \partial|\psi(\varphi)\rangle/\partial\varphi$ .

Property (i) is due to the fact that the SLD for  $p\rho(\varphi) \oplus (1-p)\sigma(\varphi)$  is a direct product of SLDs for  $p\rho(\varphi)$  and  $(1-p)\sigma(\varphi)$ , respectively. Property (ii) is a consequence of the fact that by (i) the right-hand side of (ii) can be viewed as the quantum Fisher information of the state  $p\rho(\varphi) \otimes |0\rangle\langle 0| + (1-p)\sigma(\varphi) \otimes |1\rangle\langle 1|$ , where  $|0\rangle$ ,  $|1\rangle$  are orthogonal ancillary states, while the left-hand side is  $F_Q$  of the state after tracing out the ancillary system. Furthermore,  $F_Q$  is nonincreasing under stochastic operations [27] (tracing out the ancilla is an example). Property (iii) is a consequence of Eqs. (8) and (11), since for a pure state

$$A = 2(|\psi(\varphi)\rangle\langle\psi'(\varphi)| + |\psi'(\varphi)\rangle\langle\psi(\varphi)|). \quad (15)$$

The measurement saturating the quantum Cramér-Rao bound in this case is a von Neumann measurement projecting on any orthonormal basis containing two vectors

$$|e_{\pm}\rangle = \frac{1}{\sqrt{2}}(|\psi(\varphi)\rangle \pm |\psi'(\varphi)\rangle), \quad (16)$$

where

$$|\psi'(\varphi)\rangle = \frac{1}{\mathcal{N}}[|\psi'(\varphi)\rangle - \langle\psi(\varphi)|\psi'(\varphi)\rangle|\psi(\varphi)\rangle] \quad (17)$$

is the normalized vector orthogonal to  $|\psi(\varphi)\rangle$  lying in the space spanned by  $|\psi(\varphi)\rangle$  and  $|\psi'(\varphi)\rangle$ .

#### IV. INTERFEROMETRY WITH LOSSES

Assuming that we have  $N$  photons at our disposal, we aim to find the input state that allows performing phase estimation with the best precision possible, i.e., yielding the highest value of the quantum Fisher information  $F_Q$ . In particular we consider the most general pure two-mode input state with definite photon number  $N$ ,

$$|\psi_{in}\rangle = \sum_{k=0}^N \alpha_k |k, N-k\rangle, \quad (18)$$

where  $|k, N-k\rangle$  abbreviates the Fock state  $|k\rangle_a |N-k\rangle_b$ . This class of states includes the  $N00N$  state which, in the absence of losses, leads to Heisenberg limited precision, but is very fragile in the presence of noise. We are therefore looking for states which lead possibly to a lower precision than the  $N00N$  state, but which are more robust with respect to photon losses. Moreover, although states of the form (18) seem to be a restriction, we show in Sec. VII that our treatment effectively includes states with indefinite photon number.

In the following subsections we show how states of the form (18) are influenced by photon losses, calculate its quantum Fisher information and show that the latter can be maximized (thus minimizing  $\delta\varphi$ ) by means of convex optimization methods.

##### A. Noise model

Losses are modeled by fictitious beam splitters of transmissivity  $\eta_a$ ,  $\eta_b$  in channels  $a$  and  $b$ , respectively, and cause a Fock state  $|k, N-k\rangle$  to evolve into

$$|k, N-k\rangle \mapsto \sum_{l_a=0}^k \sum_{l_b=0}^{N-k} \sqrt{B_{l_a l_b}^k} |k-l_a, N-k-l_b\rangle \otimes |l_a, l_b\rangle, \quad (19)$$

where  $|l_a, l_b\rangle$  represents the state of two ancillary modes carrying  $l_a$  and  $l_b$  photons lost from modes  $a$  and  $b$ , respectively, while

$$B_{l_a l_b}^k = \binom{k}{l_a} \binom{N-k}{l_b} \eta_a^{l_a} (\eta_a^{-1} - 1)^{l_a} \eta_b^{N-k-l_b} (\eta_b^{-1} - 1)^{l_b}. \quad (20)$$

Including the phase accumulation  $|k, N-k\rangle \mapsto e^{ik\varphi} |k, N-k\rangle$  and tracing out the ancillary modes results in the output density matrix

$$\rho(\varphi) = \sum_{l_a=0}^N \sum_{l_b=0}^{N-l_a} p_{l_a l_b} |\xi_{l_a l_b}(\varphi)\rangle \langle \xi_{l_a l_b}(\varphi)|, \quad (21)$$

where

$$|\xi_{l_a l_b}(\varphi)\rangle = \frac{1}{\sqrt{p_{l_a l_b}}} \sum_{k=l_a}^{N-l_b} \alpha_k e^{ik\varphi} \sqrt{B_{l_a l_b}^k} |k-l_a, N-k-l_b\rangle \quad (22)$$

is the conditional pure state corresponding to the event when  $l_a$  and  $l_b$  photons are lost in modes  $a$  and  $b$ , respectively, and  $p_{l_a l_b}$  is the normalization factor corresponding to the probability of that event.

Equivalently, the loss process can be described by a master equation for two independently damped harmonic oscillators with loss rates  $\gamma_{a,b} = |\ln \eta_{a,b}|/t$ , where  $t$  is time, the solution of which is given by

$$\rho = \sum_{k,l=0}^{\infty} K_{l,a} K_{k,b} \rho_{in} K_{k,b}^\dagger K_{l,a}^\dagger, \quad (23)$$

with Kraus operators

$$K_{l,a} = (1 - \eta_a)^{1/2} \eta_a^{1/2 \hat{a}^\dagger \hat{a}} \hat{a}^l / \sqrt{l!}, \quad (24)$$

where  $\hat{a}$  is the annihilation operator for mode  $a$ , and analogously for mode  $b$ . This state acquires a phase through the transformation  $\rho(\varphi) = e^{-i\varphi \hat{a}^\dagger \hat{a}} \rho e^{i\varphi \hat{a}^\dagger \hat{a}}$ . Notice that thanks to the relation

$$\hat{a}^l e^{-i\varphi \hat{a}^\dagger \hat{a}} = e^{-i\varphi \hat{a}^\dagger \hat{a}} \hat{a}^l e^{-i\varphi l} \quad (25)$$

we can commute the phase operator with the Kraus operators since the phase terms  $e^{-i\varphi l}$  cancels out. It is therefore irrelevant if photons are lost before, during or after channel  $a$  acquires its relative phase with respect to  $b$ .

### B. Calculating the Fisher information

Using Eq. (21) for the output state, one can calculate  $F_Q$  with the help of Eqs. (8) and (11). This requires diagonalization of  $\rho(\varphi)$  which can be carried out in the case of one-arm losses. In the more general case of losses in both arms an analytic calculation of  $F_Q$  turns out to be infeasible. Nevertheless, we are able to determine an upper bound to the quan-

tum Fisher information which, although not strictly tight for general input states, is very close to  $F_Q$  for the states we consider in Sec. V.

#### 1. Losses in one arm

We consider first the case  $\eta_a = \eta$ ,  $\eta_b = 1$ , i.e., when losses are present in only one arm. As can be seen from Eq. (21), in this case only states  $|\xi_{l_a l_b}(\varphi)\rangle$  with  $l_b = 0$  contribute to  $\rho(\varphi)$ . Moreover, we have  $\langle \xi_{l_0}(\varphi) | \xi_{l'_0}(\varphi) \rangle = \delta_{ll'}$ , hence we can write the output state as a direct sum

$$\rho(\varphi) = \bigoplus_{l=0}^N p_{l0} |\xi_{l0}(\varphi)\rangle \langle \xi_{l0}(\varphi)|. \quad (26)$$

Making use of Eqs. (12) and (14) we get a formula for  $F_Q$  with explicit dependence on the input state parameters  $\alpha_k$ ,

$$F_Q = 4 \left( \sum_{k=0}^N k^2 x_k - \frac{\sum_{k=0}^N x_k k B_{l_0}^k}{\sum_{k=0}^N x_k B_{l_0}^k} \right)^2, \quad (27)$$

where  $x_k = |\alpha_k|^2$ . In a more compact way the above formula can be rewritten as

$$F_Q = 2 \sum_{l=0}^N \frac{\mathbf{x}^T \cdot \mathbf{R}^{(l0)} \mathbf{x}}{\mathbf{x}^T \cdot \mathbf{b}^{(l0)}}, \quad (28)$$

where  $\mathbf{x}$  is a vector containing variables  $x_k$ , while the elements of the vector  $\mathbf{b}^{(l_a l_b)}$  and the matrix  $\mathbf{R}^{(l_a l_b)}$  are given by

$$b_k^{(l_a l_b)} = \begin{cases} B_{l_a l_b}^k & \text{if } l_a \leq k \leq N - l_b \\ 0 & \text{otherwise} \end{cases}, \quad (29)$$

$$R_{k,k'}^{(l_a l_b)} = b_k^{(l_a l_b)} (k - k')^2 b_{k'}^{(l_a l_b)}. \quad (30)$$

#### 2. Losses in two arms

If losses are present in both arms, then, in the most general case, all  $|\xi_{l_a l_b}(\varphi)\rangle$  contribute to  $\rho(\varphi)$ . States with different total number of lost photons,  $l = l_a + l_b$ , are still orthogonal. Using Eq. (12) we can therefore write

$$F_Q = \sum_{l=0}^N F_Q \left[ \sum_{l_a=0}^l p_{l_a l-l_a} |\xi_{l_a l-l_a}(\varphi)\rangle \langle \xi_{l_a l-l_a}(\varphi)| \right], \quad (31)$$

where  $F_Q[\cdot]$  denotes the quantum Fisher information of the state in brackets. Notice that states with the same  $l$  are not necessarily orthogonal. Consequently, the calculation of the above expression requires solving an eigenvalue problem which is not feasible analytically. Nevertheless, using the convexity of  $F_Q$ , Eq. (13), we obtain a bound

$$F_Q \leq \tilde{F}_Q = \sum_{l_a=0}^N \sum_{l_b=0}^{N-l_a} p_{l_a l_b} F_Q[|\xi_{l_a l_b}(\varphi)\rangle \langle \xi_{l_a l_b}(\varphi)|], \quad (32)$$

which can be calculated explicitly using Eq. (14). It is given by

$$\tilde{F}_Q = 4 \left( \sum_{k=0}^N k^2 x_k - \sum_{l=0}^N \sum_{m=0}^{N-l} \frac{\left( \sum_{k=l}^{N-m} x_k k B_{lm}^k \right)^2}{\sum_{k=l}^{N-m} x_k B_{lm}^k} \right), \quad (33)$$

or in a compact form,

$$\tilde{F}_Q = 2 \sum_{l_a=0}^N \sum_{l_b=0}^{N-l_a} \frac{\mathbf{x}^T \cdot \mathbf{R}^{(l_a l_b)} \mathbf{x}}{\mathbf{x}^T \cdot \mathbf{b}^{(l_a l_b)}}. \quad (34)$$

When losses are present only in one arm (see previous paragraph)  $F_Q = \tilde{F}_Q$ . When losses are present in both arms, however, the bound is not always tight. The difference  $\tilde{F}_Q - F_Q$  originates from the nonorthogonality of  $|\xi_{l_a l_b}(\varphi)\rangle$  for a fixed  $l = l_a + l_b$ , and physically corresponds to lack of knowledge about how many photons were lost from a particular mode. If this knowledge is not relevant then  $\tilde{F}_Q = F_Q$ . This happens, e.g., in the case of the  $N00N$  state  $\alpha_0|N, 0\rangle + \alpha_N|0, N\rangle$ , when loss of even a single photon renders the output states use-

less for phase estimation, hence the knowledge of which mode the photons were lost does not influence the value of  $F_Q$ .

### C. Concavity of Fisher information

Even with the explicit formulae for  $F_Q$  and  $\tilde{F}_Q$  given by Eqs. (28) and (34) it is in general not possible to find an analytic solution for the optimal state, i.e., values  $x_k \geq 0$ ,  $\sum_k x_k = 1$  that maximize  $F_Q$  (or  $\tilde{F}_Q$ ). However, we prove below that  $\tilde{F}_Q$  is a concave function of the  $\{x_k\}$ . Consequently, the problem amounts to the maximization of a concave function  $\tilde{F}_Q$  on a convex set. This allows for a feasible numerical constrained optimization using interior-point method routines (e.g., implemented in MATHEMATICA 6.0), and more importantly, any local maximum found is automatically the global maximum.

We prove concavity by showing that the Hessian  $H_{ij} = \frac{\partial^2 \tilde{F}_Q}{\partial x_i \partial x_j}$  is negative semidefinite, i.e., for every vector  $\mathbf{y}$  we have  $\mathbf{y} \cdot \mathbf{H} \mathbf{y} \leq 0$ . Using Eq. (34) the Hessian  $H_{ij}$  ( $i, j \in \{0, \dots, N\}$ ) reads

$$H_{ij} = 4 \sum_{l_a=0}^N \sum_{l_b=0}^{N-l_a} \frac{b_i^{(l_a l_b)} b_j^{(l_a l_b)} (\mathbf{x} \cdot \mathbf{R}^{(l_a l_b)} \mathbf{x}) - \mathbf{x} \cdot \mathbf{b}^{(l_a l_b)} [b_i^{(l_a l_b)} (\mathbf{R}^{(l_a l_b)} \mathbf{x})_j + b_j^{(l_a l_b)} (\mathbf{R}^{(l_a l_b)} \mathbf{x})_i] + \mathbf{x} \cdot \mathbf{b}^{(l_a l_b)} (\mathbf{R}^{(l_a l_b)})_{i,j} \mathbf{x} \cdot \mathbf{b}^{(l_a l_b)}}{(\mathbf{x} \cdot \mathbf{b}^{(l_a l_b)})^3}. \quad (35)$$

Since the denominator is always positive, it is sufficient to prove that for every  $l_a, l_b$  and for every vector  $\mathbf{y}$  we have

$$(\mathbf{y} \cdot \mathbf{b}^{(l_a l_b)})^2 (\mathbf{x} \cdot \mathbf{R}^{(l_a l_b)} \mathbf{x}) - (\mathbf{x} \cdot \mathbf{b}^{(l_a l_b)}) (\mathbf{y} \cdot \mathbf{b}^{(l_a l_b)}) (\mathbf{y} \cdot \mathbf{R}^{(l_a l_b)} \mathbf{x} + \mathbf{x} \cdot \mathbf{R}^{(l_a l_b)} \mathbf{y}) + (\mathbf{x} \cdot \mathbf{b}^{(l_a l_b)})^2 \mathbf{y} \cdot \mathbf{R}^{(l_a l_b)} \mathbf{y} \leq 0. \quad (36)$$

Introducing  $K_{ij} = (i-j)^2$ ,  $w_i = y_i b_i^{(l_a l_b)} (\mathbf{x} \cdot \mathbf{b}^{(l_a l_b)}) - x_i b_i^{(l_a l_b)} (\mathbf{y} \cdot \mathbf{b}^{(l_a l_b)})$  the above condition can be written equivalently as

$$\mathbf{w} \cdot \mathbf{K} \mathbf{w} \leq 0. \quad (37)$$

Since  $\sum_i w_i = 0$ , it is sufficient to prove that  $\mathbf{K}$  is negative semidefinite on the set of vectors with coefficients summing up to 0. To this end we define vectors  $\mathbf{e}^{(i)}$  ( $i=1, \dots, N$ ), where  $e_0^{(i)} = 1$ ,  $e_i^{(i)} = -1$ ,  $e_j^{(i)} = 0$  (for  $j \neq 0, j \neq i$ ), which span the space of all vectors with coefficients summing up to zero. Writing  $\mathbf{w} = \sum_{i=1}^N \beta_i \mathbf{e}^{(i)}$  and noticing that  $\mathbf{e}^{(i)} \cdot \mathbf{K} \mathbf{e}^{(j)} = -2ij$  we arrive at

$$\mathbf{w} \cdot \mathbf{K} \mathbf{w} = -2 \sum_{i,j} \beta_i \beta_j ij = -2 \left( \sum_i \beta_i i \right)^2 \leq 0, \quad (38)$$

which proves that  $\tilde{F}_Q$  is a concave function of the  $\{x_k\}$ .

## V. OPTIMAL STATES FOR PHASE ESTIMATION

In this section we discuss optimization results based on Eqs. (27) and (33) derived in the previous sections. The

quantity we analyze is  $\delta\varphi_{\min} \equiv 1/\sqrt{F_Q}$  corresponding to the best possible precision for a fixed number of measurements  $\nu$  [cf. Eq. (8)]. The only exception to this definition is in the case of the optimal state for losses in both arms where we set  $\delta\varphi_{\min} \equiv 1/\sqrt{\tilde{F}_Q}$  (see Sec. V B). We compare the optimal precision to the precision which is obtainable using various alternative states and strategies. In particular, we define the standard interferometric limit (SIL) corresponding to the precision which can be achieved in a classical reference experiment. This serves as a benchmark by which we can judge the advantage of using quantum states of light over classical ones. For a given photon number  $N$  the SIL is the precision of phase estimation using a Mach-Zehnder interferometer in which one arm gathers a phase  $\varphi$ , fed at one input port with a coherent state  $|\alpha\rangle$ , where  $|\alpha|^2 = N$ , and the vacuum at the other port. The reflectivity of the two beam splitters can be adjusted to achieve the best precision, while the measurement consists of photon counting at the two output ports. Without any additional reference beams the input coherent state should be regarded as a state with unknown phase, and effectively described as a mixture of Fock states with Poissonian statistics (see [28], and the discussion in Sec. VII A). By Eq. (12) the quantum Fisher information will be a weighted sum of quantum Fisher information calculated for each input Fock state after it passes through the first beam splitter. Taking the optimal value of transmissivity of the first beam splitter such that the ratio of the intensities in the arms

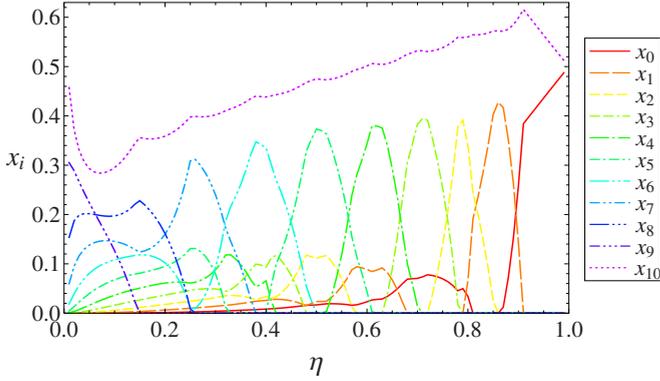


FIG. 2. (Color online) Parameters of the optimal state:  $|\psi\rangle = \sum_{i=0}^N \sqrt{x_i} |i, N-i\rangle$ , for phase estimation with  $N=10$  photons for the case of losses in one only arm, i.e.,  $\eta_a = \eta$ ,  $\eta_b = 1$ .

$a$  and  $b$  is  $\sqrt{\eta_b/\eta_a}$  the final uncertainty in phase estimation achievable with this classical strategy is given by

$$\delta\varphi_{\min, \text{SIL}} = \frac{\sqrt{\eta_a} + \sqrt{\eta_b}}{2\sqrt{N\eta_a\eta_b}}. \quad (39)$$

Note that the SIL scales in the same way as the SQL, particularly for equal losses in both arms we obtain  $\delta\varphi_{\text{SIL}} = 1/\sqrt{N\eta}$ , where  $\eta = \eta_a = \eta_b$ .

As a reference let us also calculate the minimum uncertainty achievable using  $N$  photons prepared in an unbalanced  $N00N$  state,

$$\sqrt{x_N}|N0\rangle + \sqrt{x_0}|0N\rangle. \quad (40)$$

The quantum Cramér-Rao bound yields

$$\delta\varphi_{\min, N00N} = \frac{\eta_a^{N/2} + \eta_b^{N/2}}{2N\eta_a^{N/2}\eta_b^{N/2}}, \quad (41)$$

where the optimal amplitudes are given by

$$x_0 = \frac{\eta_a^{N/2}}{\eta_a^{N/2} + \eta_b^{N/2}} \quad (42)$$

and  $x_N = 1 - x_0$ . Note that putting  $N=1$  in Eq. (41) coincides with Eq. (39). This implies that a coherent state performs equally well in phase estimation as the equivalent number of single photons sent one by one. In the following we will concentrate on the two important scenarios of losses only in channel  $a$  and equal losses in both channels.

### A. Losses in one arm

The scenario of losses in only one arm of the interferometer, i.e.,  $\eta_a = \eta$  and  $\eta_b = 1$ , is relevant, for example, if losses are induced by a sample itself. Figure 2 shows the parameters  $x_k$  of the optimal state for  $N=10$  photons as a function of  $\eta$ . The corresponding precision is shown in Fig. 3. The shaded, gray area in this figure is bounded by the SIL and the Heisenberg limit  $1/N$ , i.e., whenever a line is in this region there is an improvement over classical interferometry. The precision obtainable with the  $N00N$  state is worse than the SIL except for relatively low losses.

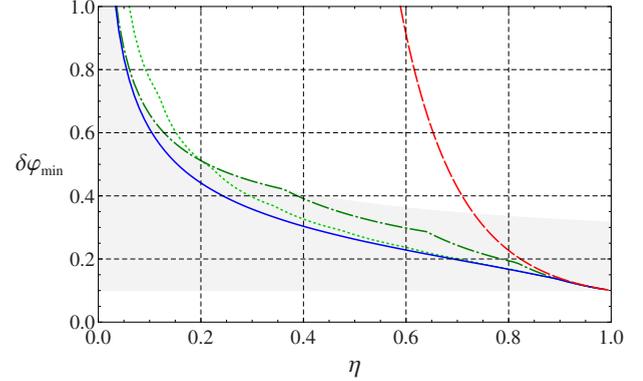


FIG. 3. (Color online) The minimal phase uncertainty achieved with various  $N=10$  photon states for losses in one arm, i.e.,  $\eta_a = \eta$ ,  $\eta_b = 1$ . Red, dashed line:  $N00N$  state; blue, solid line: optimal state (cf. Fig. 2); bright green, dotted line: optimal two-component state; dark green, dashed-dotted line:  $N00N$  chopping strategy. The shaded area is bounded by the SIL achievable with classical states and the Heisenberg limit  $1/N$ .

For transmissivities exceeding a certain threshold  $\eta > \bar{\eta}$  the optimal state is equal to the  $N00N$  state [Eq. (40)]. Note that for  $\eta=1$  the  $N00N$  state is “balanced,” i.e.,  $x_0 = x_N = 1/2$ . We will give an estimate for the threshold  $\bar{\eta}$  at the end of this subsection. For  $\eta < \bar{\eta}$  the optimal state consists of more components. As seen in Fig. 3 over a wide range we can achieve almost the same precision by using two-component states of the form

$$|\psi_{in}\rangle = \sqrt{x_k}|k, N-k\rangle + \sqrt{x_N}|N, 0\rangle, \quad (43)$$

where we numerically optimize  $k$  and the amplitude  $x_k$ . With decreasing  $\eta$  the optimal choice for  $k$  increases. This state is more robust in the presence of losses than a  $N00N$  state, since a loss of up to  $k$  photons in the first mode does not destroy the superposition, while for the  $N00N$  state a loss of even a single photon renders the state useless for phase estimation. On the other hand, a large  $N-k$  increases the sensitivity of the state with respect to an induced phase in channel  $a$ . The properties of large  $k$  and large  $N-k$  are therefore competing and the optimal result represents a trade off between phase sensitivity and robustness.

An alternative strategy that leads to an improvement over the SIL, but uses states of a simpler structure than the optimal state is the  $N00N$  “chopping” strategy which was introduced in Ref. [14]. Instead of a single  $N00N$  state, we use the same number of photons, but send them successively in  $N/n$  smaller portions using  $n$ -photon  $n00n$  states. Repeating an experiment  $N/n$  times corresponds to an  $N/n$  fold increase in Fisher information, as shown in Eq. (7). Using Eq. (41) we find that the Fisher information for the  $N00N$  chopping strategy reads

$$F_Q = \frac{N}{n} \frac{4n^2\eta^n}{(1 + \eta^{n/2})^2}. \quad (44)$$

Treating both  $n$  and  $N/n$  as real numbers with  $n \in [1, N]$ , maximization of this expression over  $n$  yields

$$\delta\varphi_{\min} = \begin{cases} \frac{1 + \sqrt{\eta}}{2\sqrt{N\eta}}; & \eta \leq \eta_0 \approx 0.228 \\ \frac{1 + \sqrt{\eta_0}}{2\sqrt{N\eta_0}} \sqrt{\frac{\ln \eta}{\ln \eta_0}}; & \eta_0 < \eta \leq \eta_0^{1/N} \\ \frac{1 + \eta^{N/2}}{2N\eta^{N/2}}; & \eta > \eta_0^{1/N}, \end{cases} \quad (45)$$

where the optimal choice of  $n$  for the three regimes indicated above is  $n=1$ ,  $n=1.478/|\ln \eta|$  (i.e., the solution of  $1 + \eta^{n/2} + n \ln \eta = 0$ ) and  $n=N$ . We see that for very small transmissivities  $\eta \leq \eta_0$  the precision is equal to the SIL, specified in Eq. (39). For higher transmissivities the chopping strategy beats the SIL, although only by a constant factor rather than in terms of scaling. For very high transmissivities  $\eta > \eta_0^{1/N}$  the best strategy is to use unchopped  $N00N$  states. We note that a similar strategy for multipartite qubit states has been devised in [11].

As mentioned above the  $N00N$  state ceases to be optimal below a threshold  $\bar{\eta}$  in which case the coefficient  $x_1$  obtains a nonzero value. By adding an infinitesimal change  $\delta x_1$  to  $x_1$  at the expense of  $x_0, x_N$  (which we assume are kept in the proportion which is optimal in the absence of  $x_1$ ), we can calculate the corresponding change in the quantum Fisher information. The change of  $x_1$  by  $\delta x_1$  increases  $F_Q$  by  $\frac{dF_Q}{dx_1} \delta x_1$ , but at the same time due to decreasing weight of the other coefficients decreases it by  $F_Q \delta x_1$ . On the whole the change of  $F_Q$  reads

$$\delta F_Q = \left( \frac{dF_Q}{dx_1} - F_Q \right) \delta x_1. \quad (46)$$

Substituting  $F_Q$  from Eq. (28) at  $x_0 = \eta^{N/2}/(1 + \eta^{N/2})$ ,  $x_N = 1 - x_0$ ,  $x_i = 0$  ( $i=2, \dots, N-1$ ) and calculating  $\delta F_Q$  we determine the value  $\bar{\eta}$  below which  $\delta F_Q / \delta x_1$  is positive. This implies that for  $\eta < \bar{\eta}$  an increase in  $x_1$  will increase  $F_Q$ . Writing  $\delta F_Q / \delta x_1$  explicitly we get

$$\frac{\delta F_Q}{\delta x_1} = \frac{4}{(1 + \eta^{N/2})^2} \left[ (1 - 2N)\eta^N - 2N(N-1)\eta^{N/2+1} + 2(N-1)^2\eta^{N/2} - N(N-2)\eta + (N-1)^2 \right]. \quad (47)$$

The roots of the expression in square brackets can be found numerically. There is one real root in the interval  $[0, 1]$  which corresponds to the threshold  $\bar{\eta}$ . The roots behave in very good approximation like  $\bar{\eta} = a^{-1/N}$  where  $a \approx 2.61$  which is obtained by a fit to the roots between  $N=5$  and  $N=100$ . For example, for  $N=10$  we find that the threshold at which the  $N00N$  state ceases to be optimal is at  $\bar{\eta} = 0.91$ , which agrees with the numerical results obtained before (see Fig. 2).

### B. Equal losses in both arms

In order to determine the optimal state in the case of equal losses in both arms of the interferometer, i.e.,  $\eta_a = \eta_b \equiv \eta$ , we use  $\tilde{F}_Q$  as given by Eq. (33) as a basis for our numerical optimization. It simplifies maximization drastically since it is

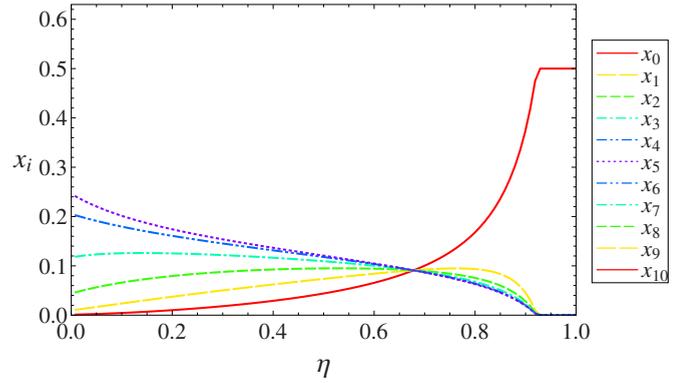


FIG. 4. (Color online) Parameters of the optimal state:  $|\psi\rangle = \sum_{i=0}^N \sqrt{x_i} |i, N-i\rangle$ , for phase estimation with  $N=10$  photons for the case of equal losses in both arms, i.e.,  $\eta_a = \eta_b = \eta$ . Due to this symmetry we have  $x_k = x_{N-k}$ .

a concave function of the  $\{x_k\}$ . As pointed out in Sec. II,  $\tilde{F}_Q$  is only an upper bound to the “true” quantum Fisher information  $F_Q$  and therefore cannot be reached by any measurement strategy if  $F_Q$  is strictly smaller than  $\tilde{F}_Q$ . However, using  $\tilde{F}_Q$  can be still very useful: we maximize  $\tilde{F}_Q$  and use the resulting input state (say,  $|\tilde{\psi}_{in}\rangle$ ) to calculate the corresponding  $F_Q(|\tilde{\psi}_{in}\rangle)$ . On the one hand, the true maximum of  $F_Q(|\psi_{in}\rangle)$  (corresponding to the truly optimal state,  $|\psi_{in}\rangle$ ) is certainly greater or equal to  $F_Q(|\tilde{\psi}_{in}\rangle)$ . On the other hand we have  $F_Q(|\psi_{in}\rangle) \leq \tilde{F}_Q(|\psi_{in}\rangle)$  due to convexity [see Eq. (32)]. But since  $\tilde{F}_Q(|\tilde{\psi}_{in}\rangle)$  is the global maximum we have  $\tilde{F}_Q(|\psi_{in}\rangle) \leq \tilde{F}_Q(|\tilde{\psi}_{in}\rangle)$ . Consequently, the true maximum must lie in between  $F_Q(|\tilde{\psi}_{in}\rangle)$  and  $\tilde{F}_Q(|\tilde{\psi}_{in}\rangle)$ . Hence, if  $|\tilde{F}_Q(|\tilde{\psi}_{in}\rangle) - F_Q(|\tilde{\psi}_{in}\rangle)|$  is small on a scale which is for this problem naturally given by the difference between the SIL and the Heisenberg limit, we obtain a very good approximation of the true maximal value of the quantum Fisher information. We showed numerically that for the examples discussed in this section this difference never exceeds 0.4%.

The result of optimization is shown in Fig. 4 for  $N=10$ . Due to the symmetry of the problem,  $\eta_a = \eta_b$ , the solution has the property  $x_k = x_{N-k}$ . For small losses the optimal state is the balanced  $N00N$  state, i.e.,  $x_0 = x_N = 1/2$ . However, below a threshold  $\bar{\eta}$  all other terms  $x_k$  obtain nonvanishing values. With decreasing  $\eta$ ,  $x_0$  and  $x_N$  get smaller, and eventually the coefficients  $x_k$  with  $k \approx N/2$  are dominant. Interestingly, this structure is qualitatively similar to a two-mode generalization of the state which is optimal for one-mode lossless phase estimation in the global approach [29]

$$|\psi_{in}\rangle = \sqrt{2/(N+2)} \sum_{k=0}^N \sin[\pi(k+1)/(N+2)] |k, N-k\rangle. \quad (48)$$

It was noted in Ref. [30] that these states perform quite well in the presence of loss, and hence could be useful for the problem we consider.

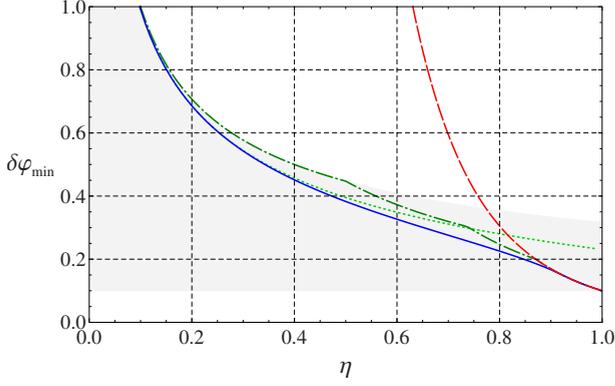


FIG. 5. (Color online) The minimal phase uncertainty achieved with various  $N=10$  photon states for equal losses in both arms, i.e.,  $\eta_a = \eta_b = \eta$ . Red, dashed line:  $N00N$  state; blue, solid line: optimal state (cf. Fig. 4); bright green, dotted line: the state optimal for lossless phase estimation in the global approach [see Eq. (48)]; dark green, dashed-dotted line: the  $N00N$  chopping strategy. The shaded area is bounded by the SIL achievable with classical states and the Heisenberg limit  $1/N$ .

The threshold  $\bar{\eta}$  can be obtained in a similar fashion as in the previous subsection. By adding an infinitesimal change  $\delta x_k$  to a component  $x_k$  other than  $x_0$  and  $x_N$  we can derive a polynomial which has a root in  $[0, 1]$  defining  $\bar{\eta}(N)$ . Since the polynomial is rather cumbersome we do not quote it here, however, the relevant root is again given in very good approximation by  $\bar{\eta} = a^{-1/N}$  where  $a \approx 2.24$  is obtained by a fit to the roots between  $N=2$  and  $N=100$ . For example for  $N=10$  (corresponding to Figs. 4 and 5) we get  $\bar{\eta} \approx 0.92$ .

Analogous to the case of losses in only one arm we can examine a  $N00N$  chopping strategy by sending “smaller”  $n00n$  states through the interferometer. The corresponding precision reads

$$\delta\varphi_{\min} = \begin{cases} \frac{1}{\sqrt{N\eta}}; & \eta \leq e^{-1} \\ \sqrt{\frac{e|\ln \eta|}{N}}; & e^{-1} < \eta \leq e^{-1/N} \\ \frac{1}{N\eta^{N/2}}; & \eta > e^{-1/N} \end{cases}. \quad (49)$$

Figure 4 shows the estimation precision for the optimal state (blue, solid line; here we use  $\delta\varphi_{\min} = 1/\sqrt{\tilde{F}_Q}$ ), the  $N00N$  state (red, dashed line), the  $N00N$  chopping strategy (dark green, dashed-dotted line), and the state (48) (bright green, dotted line). In the latter case the precision is indeed quite high in the regime of high losses, outperforming the  $N00N$  chopping strategy but for small losses it is significantly worse than both the optimal and the  $N00N$  chopping strategies [31].

## VI. OPTIMAL MEASUREMENT

As discussed in Sec. III, a measurement saturating the quantum Cramér-Rao bound always exists. For the cases

when the states  $|\xi_{l_a l_b}(\varphi)\rangle$  introduced in Sec. IV are orthogonal for different  $l=l_a+l_b$  (i.e., when  $F_Q = \tilde{F}_Q$ ) we can derive an explicit POVM saturating the bound. To this end we first assume that we perform a measurement determining the total number of photons lost,  $l$ , which projects the system onto a particular pure state  $|\xi_{l_a l_b}(\varphi)\rangle$  [32]. Subsequently, we perform the measurement saturating the Cramér-Rao bound on pure states (see Sec. III), i.e., a projection on a basis containing the vectors

$$\begin{aligned} |e_{\pm}^{l_a l_b}\rangle &= \frac{1}{\sqrt{2}} [|\xi_{l_a l_b}(\varphi)\rangle \pm |\xi_{l_a l_b}'(\varphi)\rangle] \\ &= \frac{1}{\sqrt{2}} \sum_{k=l_a}^{N-l_b} \alpha_k \sqrt{B_{l_a l_b}^k} \left[ 1 \pm \frac{i(k-l_a - \langle \hat{n}_a \rangle_{l_a l_b})}{\sqrt{(\Delta \hat{n}_a)_{l_a l_b}^2}} \right] \\ &\quad \times e^{ik\varphi} |k-l_a, N-k-l_b\rangle, \end{aligned} \quad (50)$$

where  $\langle \hat{n}_a \rangle_{l_a l_b}$ ,  $(\Delta \hat{n}_a)_{l_a l_b}^2$  are the mean photon number and the variance of the photon number in mode  $a$  for the state  $|\xi_{l_a l_b}(\varphi)\rangle$ . The above POVM depends, in general, on the phase  $\varphi$ .

There are cases in which a single POVM saturating the Cramér-Rao bound for all values of  $\varphi$  can be found. In a single photon case ( $N=1$ ), states of the form  $1/\sqrt{2}(|0,1\rangle + e^{i\chi}|1,0\rangle)$ , allow for a  $\varphi$ -independent POVM, whereas states with unbalanced weights of  $|01\rangle$ , and  $|10\rangle$  terms lead to optimal POVMs which are necessarily  $\varphi$  dependent. Moreover, for an  $N$  photon state  $|\psi\rangle = \sum_{k=0}^N \alpha_k |k, N-k\rangle$ , a single,  $\varphi$ -independent POVM saturating the Cramér-Rao bound can be found provided that the state enjoys the path-symmetry property:  $\alpha_k = \alpha_{N-k}^* e^{i\chi}$ , where  $\chi$  is a fixed phase the same for all components [33]. In general, due to losses the conditional pure states that appear in our problem lack the above mentioned symmetry and consequently the optimal POVM varies with the change of  $\varphi$ .

## VII. MORE GENERAL INPUT STATES

One could argue that the  $N$  photon state given by Eq. (18) does not represent the most general way of using  $N$  photons to determine the relative phase in an interferometer: instead of considering a state with a *definite* photon number  $N$ , one could consider a state which is a superposition of terms with different total photon number, and only imply a constraint that the *mean* photon number is  $N$ . Furthermore, by using states of the form (18) it is assumed that all photons in an arm of the interferometer are indistinguishable, i.e., occupy a common spatio-temporal mode. This greatly limits the class of considered states since the Hilbert space dimension of states sent into the interferometer would be of the order  $2^N$  if the photons were distinguishable, while for indistinguishable photons it is  $N+1$ . In the next two subsections we show that none of these generalizations improves the phase estimation precision, and consequently it is sufficient to consider the states of the form given by Eq. (18).

### A. States with indefinite photon number

We emphasize that in this paper we consider closed systems, i.e., there are no additional reference beams, neither

classical nor quantum, since any additional beam would contain photons and should therefore be explicitly taken into account for the determination of the required resources.

Consider a superposition of input states of the interferometer with different, but definite photon numbers  $n$ ,

$$|\Psi_{in}\rangle = \sum_{n=0}^{\infty} \beta_n |\psi_{in}^{(n)}\rangle, \quad (51)$$

where the mean total photon number is  $N = \sum_{n=0}^{\infty} |\beta_n|^2 n$ . Under free evolution a term with  $n$  photons acquires a phase  $e^{-in\omega t}$ , i.e., terms with different  $n$  evolve with different frequencies. In the absence of an additional reference beam which allows for clock synchronization between the sender and the receiver the relative phases between terms with different  $n$  become unobservable and the state given by Eq. (51) is physically equivalent to a mixture [28,34],

$$\rho_{in} = \sum_{n=0}^{\infty} |\beta_n|^2 |\psi_{in}^{(n)}\rangle\langle\psi_{in}^{(n)}|. \quad (52)$$

Moreover, since the quantum Fisher information is convex (see Sec. III), we have

$$F_Q(\rho_{in}) \leq \sum_{n=0}^{\infty} |\beta_n|^2 F_Q(|\psi_{in}^{(n)}\rangle\langle\psi_{in}^{(n)}|). \quad (53)$$

Consequently it is always better to send a state with a fixed number of photons  $|\psi_{in}^{(n)}\rangle$ , with probability  $|\beta_n|^2$ , rather than to use a superposition (which is effectively a mixture). The analysis can thus be restricted to states with definite photon number without compromising optimality.

### B. Distinguishability of photons

In this subsection we consider general  $N$  photon input states where the photons are distinguishable which is the case, e.g., if they are sent in different time bins. A state of this type can be written as

$$|\psi_{in}\rangle = \sum_{\mathbf{k}=0}^{\mathbf{1}} \alpha_{\mathbf{k}} |\mathbf{k}\rangle, \quad (54)$$

where  $\mathbf{k}$  represents a binary sequence of length  $N$ , and  $\mathbf{0} \equiv \{0, \dots, 0\}$ ,  $\mathbf{1} \equiv \{1, \dots, 1\}$ . Summation from a sequence to a sequence should be understood using partial ordering between sequences, i.e.,  $\mathbf{k} \leq \mathbf{k}' \Leftrightarrow k_i \leq k'_i \forall i$ . Therefore the summation is performed over all binary sequences. Furthermore  $|\mathbf{k}\rangle = |k_1\rangle \otimes \dots \otimes |k_N\rangle$  denotes a state of  $N$  photons, where  $k_i = 0(1)$  means that the photon sent in the  $i$ th time bin propagates in arm  $a$  ( $b$ ) of the interferometer. Mathematically, we deal here with an arbitrary  $N$ -qubit state, which lives in a  $2^N$  dimensional space. Notice that the indistinguishable case is recovered once we consider only the fully symmetric (bosonic) subspace of  $N$  qubits.

$\tilde{F}_Q$  can be calculated in a similar way as in Sec. IV B and reads

$$\tilde{F}_Q = 4 \left( \sum_{\mathbf{k}=0}^{\mathbf{1}} \bar{\mathbf{k}}^2 x_{\mathbf{k}} - \sum_{l_a=0}^{\mathbf{1}} \sum_{l_b=0}^{\mathbf{1}-l_a} \frac{\left( \sum_{k=l_a}^{1-l_b} \bar{\mathbf{k}} x_{\mathbf{k}} B_{l_a l_b}^{\mathbf{k}} \right)^2}{\sum_{k=l_a} x_{\mathbf{k}} B_{l_a l_b}^{\mathbf{k}}} \right), \quad (55)$$

where boldface symbols denote binary strings of length  $N$ ,  $x_{\mathbf{k}} = |\alpha_{\mathbf{k}}|^2$ ,  $\bar{\mathbf{k}}$  denotes the number of 1s in the sequence  $\mathbf{k}$ , and

$$B_{l_a l_b}^{\mathbf{k}} = (1 - \eta_a)^{\bar{l}_a} \eta_a^{\bar{\mathbf{k}} - \bar{l}_a} (1 - \eta_b)^{\bar{l}_b} \eta_b^{N - \bar{\mathbf{k}} - \bar{l}_b}. \quad (56)$$

Again  $\tilde{F}_Q = F_Q$  provided that knowledge from which mode a photon was lost is not relevant or self-evident, e.g., when  $\eta_b = 1$ . Let us now symmetrize the state of  $N$  photons. This corresponds to replacing  $x_{\mathbf{k}}$  with  $1/N! \sum_{\sigma} x_{\sigma(\mathbf{k})}$ , where the summation is performed over all permutations of an  $N$  element set. Notice that the first term within the parenthesis in Eq. (55) is not affected by symmetrization. Let us denote  $a_{l_a l_b} = \sum_{k=l_a}^{1-l_b} \bar{\mathbf{k}} x_{\mathbf{k}} B_{l_a l_b}^{\mathbf{k}}$  and  $b_{l_a l_b} = \sum_{k=l_a}^{1-l_b} x_{\mathbf{k}} B_{l_a l_b}^{\mathbf{k}}$ . Then the subtracted term in Eq. (55) reads  $\sum_{l_a=0}^{\mathbf{1}} \sum_{l_b=0}^{\mathbf{1}-l_a} \frac{a_{l_a l_b}^2}{b_{l_a l_b}}$ . Performing symmetrization corresponds to replacing  $a_{l_a l_b}$  and  $b_{l_a l_b}$  with  $(1/N!) \sum_{\sigma} a_{\sigma(l_a), \sigma(l_b)}$  and  $(1/N!) \sum_{\sigma} b_{\sigma(l_a), \sigma(l_b)}$ , respectively. Since  $b_{l_a l_b}$  are positive, the following inequality holds

$$\frac{a_{l_a l_b}^2}{b_{l_a l_b}} + \frac{a_{l'_a l'_b}^2}{b_{l'_a l'_b}} \geq 2 \frac{\left[ \frac{1}{2} (a_{l_a l_b} + a_{l'_a l'_b}) \right]^2}{\frac{1}{2} (b_{l_a l_b} + b_{l'_a l'_b})}, \quad (57)$$

which shows that symmetrizing can only decrease the subtracted fraction, hence increase the Fisher information. This proves that the optimal states are symmetric states living in the bosonic subspace. Consequently, using indistinguishable photons is sufficient to obtain the optimal Fisher information.

A physical intuition behind the above derivation is the following. Notice that if there are no losses, the optimal  $N$  qubit state has the form  $(1/\sqrt{2})(|\mathbf{0}\rangle + |\mathbf{1}\rangle)$ —it is the  $NOON$  state which lives in the fully symmetric subspace. In this case there is no advantage in using distinguishable photons. If there are losses, things get even worse for distinguishable photons. There is still no advantage in terms of phase sensitivity, yet when a photon is lost, the *knowledge of which photon was lost* additionally harms the quantum superposition. Hence, it is optimal to use states from a fully symmetric subspace, where it is not possible to tell which photon was actually lost.

## VIII. SUMMARY

We have analyzed the optimal way of using  $N$ -photon states for phase interferometry in the presence of losses. We have derived an explicit formula for the quantum Fisher information in the case when losses are present in one arm of the interferometer, and have provided a useful bound for the case of losses in both arms. Using the quantum Fisher information as a figure of merit, we have found the optimal states, investigated their advantage over various quantum and classical strategies. The optimal measurement saturating the Cramér-Rao bound has been presented, and it has been

proven that in general it varies depending on the phase in the vicinity of which we perform estimation.

A close inspection of the properties of the quantum Fisher information showed that it is optimal to use a quantum state with a definite number of indistinguishable photons. Neither allowing superpositions of different total photon-number terms, nor making photons distinguishable can improve estimation precision.

## ACKNOWLEDGMENTS

This research was supported by the EPSRC (U.K.) through the QIP IRC (Grant No. GR/S82176/01), the AFOSR through the EOARD, the European Commission under the Integrated Project QAP (Contract No. 015848), the Royal Society and the Polish MNiSW (Grant No. N N202 1489 33).

- 
- [1] V. Giovannetti, S. Lloyd, and L. Maccone, *Science* **306**, 1330 (2004).
- [2] V. Giovannetti, S. Lloyd, and L. Maccone, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [3] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, *Nature (London)* **429**, 161 (2004).
- [4] H. S. Eisenberg, J. F. Hodelin, G. Khoury, and D. Bouwmeester, *Phys. Rev. Lett.* **94**, 090502 (2005).
- [5] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi, *Science* **316**, 726 (2007).
- [6] K. J. Resch, K. L. Pregnell, R. Prevedel, A. Gilchrist, G. J. Pryde, J. L. O'Brien, and A. G. White, *Phys. Rev. Lett.* **98**, 223601 (2007).
- [7] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, *Nature (London)* **450**, 393 (2007).
- [8] B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde, *New J. Phys.* **11**, 073023 (2009).
- [9] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, *Phys. Rev. A* **54**, R4649 (1996).
- [10] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, *Phys. Rev. Lett.* **79**, 3865 (1997).
- [11] A. Shaji and C. M. Caves, *Phys. Rev. A* **76**, 032111 (2007).
- [12] M. Sarovar and G. J. Milburn, *J. Phys. A* **39**, 8487 (2006).
- [13] S. D. Huver, C. F. Wildfeuer, and J. P. Dowling, *Phys. Rev. A* **78**, 063828 (2008).
- [14] U. Dörner, R. Demkowicz-Dobrzanski, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, *Phys. Rev. Lett.* **102**, 040403 (2009).
- [15] M. A. Rubin and S. Kaushik, *Phys. Rev. A* **75**, 053805 (2007).
- [16] M. H. G. Gilbert and Y. Weinstein, *J. Opt. Soc. Am. B* **25**, 1336 (2008).
- [17] D. Meiser and M. J. Holland, *New J. Phys.* **10**, 073014 (2008).
- [18] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [19] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
- [20] S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [21] S. L. Braunstein, C. M. Caves, and G. J. Milburn, *Ann. Phys. (N.Y.)* **247**, 135 (1996).
- [22] Sometimes a weaker condition of *asymptotic local unbiasedness* is imposed, i.e., unbiasedness of estimation when the estimator is calculated using data coming from a very large number of repetitions of an experiment (see, e.g., [20]).
- [23] O. E. Barndorff-Nielsen and R. D. Gill, *J. Phys. A* **33**, 4481 (2000).
- [24] G. A. Durkin and J. P. Dowling, *Phys. Rev. Lett.* **99**, 070801 (2007).
- [25] R. A. Fisher, *Proc. Cambridge Philos. Soc.* **22**, 700 (1925).
- [26] A. Fujiwara, *Phys. Rev. A* **63**, 042304 (2001).
- [27] D. Petz, *Numer. Linear Algebra Appl.* **244**, 81 (1996).
- [28] K. Mølmer, *Phys. Rev. A* **55**, 3195 (1997).
- [29] D. W. Berry and H. M. Wiseman, *Phys. Rev. Lett.* **85**, 5098 (2000).
- [30] L. Maccone and G. De Cillis, *Phys. Rev. A* **79**, 023812 (2009).
- [31] Claims in [30] that the state given by Eq. (48) outperforms even the optimal states we have presented arise from taking a different perspective in quantifying the resources. See Sec. VII A for more discussion.
- [32] Up to now, no effective nondestructive photon-number measurement was realized. Nevertheless, as most quantum optical experiments work in the postselection paradigm, this is not really a limitation since the number of photons can be detected afterward and only appropriate events taken into account.
- [33] H. F. Hofmann, *Phys. Rev. A* **79**, 033822 (2009).
- [34] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, *Int. J. Quantum Inf.* **4**, 17 (2006).