Fock state interferometry for quantum enhanced phase discrimination

Reihaneh Shahrokhshahi,1 Saikat Guha,2 and Olivier Pfister3

1University of Virginia, Dept. of Physics, 382 McCormick Rd., Charlottesville, VA 22904-4714, USA
2College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA

(Dated: February 16, 2021)

We study Fock state interferometry, consisting of a Mach-Zehnder Interferometer with two Fock state inputs and photon-number-resolved detection at the two outputs. We show that it allows discrimination of a discrete number of apriori-known optical phase shifts with an error probability lower than what is feasible with classical techniques under a mean photon number constraint. We compare its performance with the optimal quantum probe for M-ary phase discrimination, which unlike our probe, is difficult to prepare. Our technique further allows discriminating a null phase shift from an increasingly small one at zero probability of error under ideal conditions, a feature impossible to attain using classical probe light. Finally, we describe one application to quantum reading with binary phase-encoded memory pixels.

I. INTRODUCTION

Quantum mechanics dictates the fundamental limit of phase estimation. A typical interferometer, such as the Michelson or Mach-Zehnder interferometer, probes a phase shift by splitting an input field into two mutually coherent fields, each probing a different path. The fields are then recombined and the resulting total intensity yields interference fringes which inform on the phase-difference between the two paths. The quantum physics of interferometry hinges on the initial splitting of the input field: by unitarity, the two split outputs necessarily call for two inputs, in of which is in a vacuum state. As shown by Carl Caves [1], it the yields the classical — a.k.a. beam-splitter shot-noise — limit

$$\Delta \theta_{cl} \sim \langle N \rangle^{-\frac{1}{2}}, \quad (1)$$

where \( \theta \) is the phase difference between the arms of an interferometer and \( \langle N \rangle \) is the total average number of photons in the interferometer. The classical limit, however, doesn’t give the ultimate phase precision, which is fixed by the Heisenberg number-phase Heisenberg inequality [2] and bounded by the Heisenberg limit (HL)

$$\Delta \theta_{HI} \sim \langle N \rangle^{-1}. \quad (2)$$

In order to reach the HL, the vacuum input must be replaced by a nonclassical state such as a squeezed state [3]. Other possibilities were proposed [4], in particular the use of correlated Fock states [5] or superpositions or statistical mixtures thereof [6]. Many theoretical studies of nonclassical inputs yielding the HL have been conducted [7], in particular applying information theory to quantum physics [8], which has led to the definition of quantum metrology [9]. In addition, a recent study has shown that ensuring phase estimation at the HL requires, in general, losses to be bounded by \( \langle N \rangle^{-1} \) [10].

In this paper, we investigate the properties of the highly nonclassical, yet conceptually simple interferometry which we call Fock-state interferometry (FSI), as depicted in Fig. 1 in all of the paper, FSI consists in a Mach-Zehnder Interferometer (MZI) with Fock state input and photon-number-resolved detectors (PNRDs). One should note that FSI was recently implemented experimentally, and the effect of photon loss was studied [11]. FSI was also studied theoretically with multimode inputs [12]. Here, we focus on FSI not for phase estimation as much as for phase discrimination.

It may be useful to give a simple insight on the expected impact of such a purely corpuscular interferometer input on purely undularory interferometric performance. It is simple to show, as we do below, that FSI benefits general phase estimation only when the input state satisfies \( n_a = n_b \) [5]. We outline a basic derivation, using the Schwinger spin representation of two boson fields [13], which we’ll use in all of the paper and which is presented, along with its application to quantum interferometry, in Appendix A. The Schwinger spin formalism allows one

\[ n_{a_{out}} \]

\[ n_{b_{out}} \]

FIG. 1. Phase discrimination by Fock-state interferometry. The Fock state \(| n_a \rangle_a | n_b \rangle_b \) is input into a Mach-Zehnder interferometer of phase difference \( \theta \) and the interferometer’s output is measured by photon-number-resolved detectors.
to use the convenient formalism of quantum angular momentum and SO(3) rotation matrices to perform quantum optics calculations.

To begin with, the expectation value of the output photon number difference is

\[
\langle N_a - N_b \rangle_{\text{out}} = \langle J_z \rangle_{\text{out}} = 2m \cos \theta \quad (3)
\]

\[
= (n_a - n_b) \cos \theta \quad (4)
\]

which constitutes the quantum expression of an interference fringe.

Note that, when \( n_a = n_b \iff m = 0 \), the output photon number difference is zero for all values of \( \theta \) hence no direct fringe is present and other methods are required to access phase information. However, these methods yield Heisenberg-limited performance \([3, 4]\).

When \( n_a \neq n_b \iff m \neq 0 \), a direct interference signal is present but we show that direct measurement performance of FSI is, in fact, at the classical limit. First, we evaluate the quantum standard deviation of the output photon-number difference, using \( j = (n_a + n_b)/2 \) and \( m = (n_a - n_b)/2 \) for input state \( |jm\rangle_z = |n_a\rangle_a |n_b\rangle_b \),

\[
\Delta J_z^{\text{out}} = |\sin \theta| \left\{ \frac{1}{2} [j(j+1) - m^2] \right\}^{\frac{1}{2}} \quad (6)
\]

from which we obtain the theoretical phase error (whose expression is invalid for \( n_a = n_b \))

\[
\Delta \theta = \frac{\Delta J_z^{\text{out}}}{\partial J_z^{\text{out}}} \quad (7)
\]

\[
\Delta \theta = \left[ j(j+1) - \frac{1}{2} \right]^{\frac{1}{2}} \quad (8)
\]

which is only valid for \( m \neq 0 \), i.e., \( n_a \neq n_b \). Since the total photon number \( j \) is a constant here, it is easy to see that the minimum error can only be obtained by maximizing \( m^2 \), i.e., for \( m = \pm j \ (n_{b,a} = 0 \) respectively), in which case we get

\[
\Delta \theta_{\text{min}} = (2j)^{-\frac{1}{2}} \quad (9)
\]

which is the classical limit, Eq. (1), as was first shown by Caves [1].

We show in this paper that FSI can yield enhanced interferometry performance and break the classical limit in cases less general than the estimation of an unknown phase, namely the discrimination of two or more predetermined phase shifts. The rationales for this study are multiple: (i), the theoretical simplicity of the experimental setup; (ii), its translation into sophisticated experimental concepts that are nonetheless coming of age, i.e., on-demand photon sources and photon-number-resolving detectors; (iii), the availability of concrete applications of phase discrimination, such as the quantum reading of classical optical memories.

This paper is organized as follows, in Section II we study the problem of discriminating the finite number \( M \geq 2 \) of optical phase shifts for M-ary Phase Shift Keying (MPSK) [14] and show that our model can discriminate between two or three phases with respectively zero and near zero error probability using only few photons, and outperforms phase discrimination schemes using coherent states with heterodyne and homodyne receivers. In Section III we study the quantum optical reading of digital memory and show it outperforms its classical counterpart. We then conclude.

II. PHASE DISCRIMINATION

II.1. Introduction

In optical communication terms, the ability to discriminate between \( M \) optical phase shifts can be beneficial to MPSK, a digital modulation scheme that conveys \( M \) messages by modulating the optical phase of a probe signal. Here we show that using FSI we can accurately discriminate between \( M = 2, 3 \) optical phase shifts, using a single photon count. Moreover, we show that FSI allows discrimination of the information encoded in increasingly smaller phase shifts.

For simplicity, we first explain the concept with \( M = 2 \) phases and then generalize it to higher values of \( M \). We consider an unknown fixed phase \( \theta \) of MZI, which can take one of \( M = 2 \) values, denoted \( \theta_1 \) and \( \theta_2 \) and we define the estimated phase \( \hat{\theta} \). Four different scenarios can occur during the phase discrimination process. If the initial phase \( \theta = \theta_{1,2} \) and the estimated phase \( \hat{\theta} = \theta_{1,2} \), respectively, then we have success. Else \( \hat{\theta} = \theta_{1,2} \), and we have an error. A natural criterion to measure interferometer performance in the phase discrimination problem will then be the error probability, \( P_e \), which we’ll define later.

We consider a MZI with a \( |j \mu\rangle \) input and whose phase \( \theta \) can be either of two predetermined values \( \theta_{1,2} \). We then perform a single \( J_z \) measurement of the photon number difference at the output ports, of result \( \mu' \), and make a decision about the phase shift based on maximum likelihood algorithm: knowing the probability distribution \( P(\mu', \mu | \theta) \) of the interferometer (Table I), we compare both cases \( \theta = \theta_1 \) and \( \theta = \theta_2 \) for a given measurement outcome and assign the estimated phase shift \( \hat{\theta} \) to the phase which is more likely to result in this specific outcome \( \mu' \).

The algorithm is thus

\[
\text{if } P(\mu', \mu | \theta_1) \geq P(\mu', \mu | \theta_2) \quad \text{then} \quad \begin{cases} P(\hat{\theta} = \theta_1 | \theta = \theta_1) = P(\mu', \mu | \theta_1) \quad \text{— success} \\ P(\hat{\theta} = \theta_1 | \theta = \theta_2) = P(\mu', \mu | \theta_2) \quad \text{— failure} \end{cases}
\]

\[
\text{else} \quad \begin{cases} P(\hat{\theta} = \theta_2 | \theta = \theta_1) = P(\mu', \mu | \theta_1) \quad \text{— success} \\ P(\hat{\theta} = \theta_2 | \theta = \theta_2) = P(\mu', \mu | \theta_2) \quad \text{— failure} \end{cases}
\]
TABLE I. Probability distribution $P(\mu', \mu|\theta)$. The possible measurement outcomes are denoted by $\mu'$ (columns) and possible phases by $\theta_{1,2}$ (rows). Each element of this array is the probability of measuring $\mu' \in [-j, j]$, given phase $\theta$.

| $\theta$  | $\mu'$ | $-j$ | $\ldots$ | $m$ | $\ldots$ | $j$ |
|----------|--------|------|----------|-----|----------|-----|
| $\theta_1$ | $P(-j, \mu|\theta_1)$ | $\ldots$ | $P(m, \mu|\theta_1)$ | $\ldots$ | $P(j, \mu|\theta_1)$ |
| $\theta_2$ | $P(-j, \mu|\theta_2)$ | $\ldots$ | $P(m, \mu|\theta_2)$ | $\ldots$ | $P(j, \mu|\theta_2)$ |

For this procedure to be error free, one would need:

$$P(\hat{\theta} = \theta_{1,2}|\theta = \theta_{2,1}) = 0.$$ (10)

Of course, this is not the case in general, and the average error probability is given by:

$$P_e = \sum_{i,j \neq i} P(\theta_i)P(\hat{\theta} = \theta_j|\theta = \theta_i).$$ (11)

However, we now show that, for a judicious choice of phases $\theta_{1,2}$ and input $\mu$, phase discrimination with an FSI may perform better than classical methods such as a coherent state probe and homodyne or heterodyne detection.

II.2. Binary phase discrimination

The analytic expressions of probabilities are given by rotation matrix elements in the Schwinger representation, Eq. (A23), Appendix A. Without loss of generality, we may elect to set $\theta_1 = 0$ as this entails

$$P(\mu', \mu|0) = d_{\mu',\mu}^0(0)^2 = \delta_{\mu',\mu},$$ (12)

and simplifies the situation. The problem will then reduce to discriminating $\theta_2 = \theta$ against $\theta_1 = 0$. Note that this is still different from general phase estimation — again classically-limited for a FSI — as we’ll restrict $\theta$ to the values that will allow optimized performance.

Using information theory — see Appendix B for a brief review — we will evaluate two important figures of merit of phase discrimination: first, the error probability, Eq. (11), and also the mutual information $I(\theta; \hat{\theta})$ between $\theta$ and $\hat{\theta}$, Eq. (B14).

II.2.1. Influence of the total photon number

We study the behavior of FSI for input states with different total photon number $n_a + n_b = 2j$. We calculate the error probability, and also the mutual information, for all $\theta$. We then determine optimal phases for each input state, based on the performance.

We first consider binary ($M = 2$) phase discrimination for the two-photon inputs $|1\rangle_a |1\rangle_b$ and $|2\rangle_a |0\rangle_b$, which are the respective Schwinger spin states $|10\rangle_z$ and $|11\rangle_z$. Figure 2 displays the error probability versus the phase. As can be seen on the figure, both inputs yield zero error probability. However, the smallest phase $\theta$ such that $P_e(\theta) = 0$ is $\pi$ for a “classical” FSI (vacuum input, i.e., $m = \pm j$) and $\pi/2$ for an $m = 0$ input. An immediate question is therefore whether this trend continues and larger photon numbers yield discrimination of smaller phases.

The 6-photon ($j = 3$) and 8-photon ($j = 4$) cases are displayed in Fig. 3(a) and Fig. 3(b), respectively. These situations are richer because more photon number partitions are possible at the interferometer’s input. Two trends are clearly visible: (i), from Fig. 3(a) and Fig. 3(b) individually, i.e., for a given $j$, the first zero of $P_e(\theta)$, or “smallest optimum phase,” decreases as $|m|$ decreases and, (ii), from Fig. 3(a) and Fig. 3(b) together, the smallest optimum phase decreases as $j$ increases.

In order to confirm the latter, we computed the smallest optimum phase $\theta_{\text{sop}}$ for larger photon numbers in the $m = 0$ and $m = j$ cases, results are plotted in Fig. 4. For the values up to $2j = 10$ computed in Fig. 4, a fit for the balanced Fock state with $m = 0$ (red plot) yields

$$\theta_{\text{sop}} = 2.67(6) \times (2j)^{-0.77(2)},$$ (13)

where the parenthetical numbers (6) and (2) represent fit uncertainty. While we do not yet have any certainty about the theoretical form of the dependence of $\theta_{\text{sop}}$ on $j$, these results clearly hint at the ability for “twin” Fock states $|n\rangle_a |n\rangle_b = |j = n m = 0\rangle_z$ to allow the discrimination of phase shifts $j^{-3/4}$,
which beats the classical limit $j^{-1/2}$. This is, indeed, proven by our calculations up to $2j = 10$.

This result is reminiscent of Holland-Burnett interferometry, which employs the same input state. However, the measurement processing is entirely different as it uses maximum likelihood, as opposed to Bayesian reconstruction.

We now turn to another form of comparison of FSI phase discrimination with conventional methods, by comparing it with the coherent-state-based homodyne detection and Dolinar receivers, as well as with the use of phase eigenstates.

### II.2.2. Comparison with coherent state input methods and with phase eigenstates

Various other states can be used as interferometer input. Coherent states, in particular, present the advantage of being readily accessible from the output of a well stabilized laser.

In binary phase-shift-keyed encoding using coherent states, the phase-modulated coherent states are given as $|\alpha\rangle$ and $|\alpha\rangle$, i.e., two fields of opposite phase. In the low photon number regime, these states have small amplitudes ($|\alpha| = \sqrt{j}$) and are largely overlapping (due to nonorthogonality of coherent states), so the ability to successfully distinguish these states at the receiver is limited, ultimately by the Helstrom bound [8]. Various detection schemes can approach this bound for coherent state discrimination. We consider two of them: the homodyne receiver and the Dolinar receiver.

The former is the simplest possible receiver relying on Gaussian (Wigner) function operations [15]. The error probability for coherent state discrimination using a homodyne receiver is given by [16]:

$$P_e(\alpha) = \frac{1}{2} \left[ 1 - \text{erf} \left( \frac{|\alpha|}{2} \right) \right].$$

(14)

The latter is an adaptive measurement scheme, proposed by Dolinar in 1973, and which reaches the Helstrom bound for discriminating between two pure coherent states 17. The Dolinar receiver is based on a combination of photon counting and real-time feedback control. The minimum error probability using a Dolinar receiver is

$$P_e^{\text{min}} = \frac{1}{2} \left( 1 - \sqrt{1 - e^{-4|\alpha|^2}} \right),$$

(15)

which is the lowest possible error in distinguishing between two pure coherent states. Therefore, one reaches zero error probability only asymptotically, using high-amplitude coherent states.

Finally, we mention a theoretical proposal, which doesn’t rely on coherent states, by Nair et al. [18] for the quantum state inside (between the two beam-splitters of) the MZI, rather than at its input. When
of the optical phase operator \[19\]

This quantum state is Pegg and Barnett’s eigenstate which can resolve \(\pi\) and \(\pi\) radians, versus \(2\pi\) radians.

FIG. 5. Error probability for discriminating between 0 and \(\pi\) radians, versus \(j\) (where \(2j\) is the total photon number).

\[|\varphi\rangle = \frac{1}{\sqrt{m}} \sum_{n=0}^{m-1} e^{-in\varphi} |n\rangle \]

(16)

taken at \(\varphi = 0\)

\[|\varphi = 0\rangle = \frac{1}{\sqrt{m}} \sum_{n=0}^{m-1} |n\rangle \]

(17)

This state is intended to probe the phase shift in the “signal” arm of the MZI, which it can do at the Heisenberg limit. For \(m = 2\) \((j \leq 1)\), the minimum error probability is

\[P_{e} = \frac{1}{2} - \sqrt{j(1-j)}. \]

(18)

We now turn to the comparison of the error probabilities for binary phase discrimination (one of the phases being set to zero) of Fock-state interferometry, coherent-state homodyne detection, the Dolinar receiver, and phase-eigenstate probe. The respective error probabilities for discriminating between 0 and \(\pi\) radians are plotted in Fig. 5. As can be seen on the figure, FSI’s performance is remarkable as it matches that of the optimal phase-eigenstate probe in this \((0, \pi)\) case, and, like the phase-eigenstate probe, yields a clear advantage over the Dolinar receiver for small photon numbers. This is also of interest as FSI, unlike the phase-eigenstate probe, has a clear experimental implementation.

In addition, we recall that FSI offers the additional capability of resolving decreasing phase shifts as the total photon number increases (Fig. 4), as per Eq. (13).

II.3. Ternary phase discrimination

We now turn to the extension of the previous problem to discriminating three phases \((0, \theta_1, \theta_2)\) — one of them being, again, set to zero for convenience and without loss of generality. Again, the error probability \(P_e\), is a natural criterion to assess the performance of the phase discrimination. For all phases equiprobable, the error probability is, from Eq. (11),

\[P_{e} = \frac{1}{3} [P(\theta_1) + P(\theta_2) + P(\theta_1|0) + P(\theta_2|0) + P(\theta_2|\theta_1)] \]

(19)

Figure 6 displays the error probability for all possible combinations of \((0, \theta_1, \theta_2)\), for a twin-photon input.

In Fig. 7, we plot the error probability for the other two possible Fock-state inputs with \(j = 2\). As for the binary case, the discrimination performance depends very much on the input state. Also, a shared trend with the binary case is that the error probability displays more oscillations, and therefore more local minima, versus the phase angle(s) as \(m\) decreases, with the most favorable situation occurring for \(m = 0\). This case also allows discrimination of smaller phase shifts, as for the binary case.

We further determined optimum phase values for different input states up to \(j = 6\). These are listed, along with the error probability and mutual information [Eq. (16)], in Table II. Ternary discrimination does differ from binary discrimination: while the trend that balanced inputs allow the discrimination of smaller phases as \(j\) increases is confirmed by examination of all results for \(m = 0\), one can notice

\[|\alpha| = j \]
Fig. 7. Error probability ($P_e$) vs phase shifts $\theta_1$ and $\theta_2$, for optical phase discrimination between three phase shifts $(0, \theta_1, \theta_2)$. MZI input is a), $|3\rangle_a|1\rangle_b = |2, 1\rangle_z$ and b), $|4\rangle_a|0\rangle_b = |2, 2\rangle_z$.

Table II. Optimal phases $\theta_1^{opt}$, $\theta_2^{opt}$ that attain the minimum error probability $P_{e}^{min}$, and corresponding mutual information $I((\theta); \{\theta\})$ for all input states with $2 \leq j \leq 6$. (The maximum value of $I((\theta); \{\theta\})$ here is $\log_2 M = 1.585$.)

| Input state | $\theta_1^{opt}$ | $\theta_2^{opt}$ | $10^{-3} P_{e}^{min}$ | $I((\theta); \{\theta\})$ |
|-------------|-----------------|-----------------|------------------------|--------------------------|
| $|2\rangle_a$ | $\frac{\pi}{2}$ | $\frac{3\pi}{4}$ | 160 | 0.93 |
| $|21\rangle_a$ | $|3\rangle_a$ | $|1\rangle_b$ | 160 | 0.97 |
| $|22\rangle_a$ | $|4\rangle_a$ | $|0\rangle_b$ | 40 | 1.35 |
| $|30\rangle_a$ | $|3\rangle_a$ | $|3\rangle_b$ | $0.67 \frac{\pi}{4}$ | 140 | 1.13 |
| $|31\rangle_a$ | $|4\rangle_a$ | $|2\rangle_b$ | $\frac{\pi}{2}$ | 8 | 1.50 |
| $|32\rangle_a$ | $|5\rangle_a$ | $|1\rangle_b$ | 2.32 $3.16 \frac{\pi}{2}$ | 5 | 1.55 |
| $|33\rangle_a$ | $|6\rangle_a$ | $|0\rangle_b$ | $\frac{\pi}{3}$ | 48 | 1.30 |
| $|40\rangle_a$ | $|4\rangle_a$ | $|4\rangle_b$ | 0.55 $1.2 \frac{\pi}{2}$ | 120 | 1.12 |
| $|41\rangle_a$ | $|5\rangle_a$ | $|3\rangle_b$ | 1.2 $\frac{\pi}{4}$ | 5 | 1.52 |
| $|42\rangle_a$ | $|6\rangle_a$ | $|2\rangle_b$ | 0.6 $\frac{\pi}{2}$ | 2 | 1.55 |
| $|43\rangle_a$ | $|7\rangle_a$ | $|1\rangle_b$ | 0.66 $\frac{\pi}{2}$ | 3 | 1.56 |
| $|44\rangle_a$ | $|8\rangle_a$ | $|0\rangle_b$ | $\frac{\pi}{4}$ | 3 | 1.41 |
| $|50\rangle_a$ | $|5\rangle_a$ | $|5\rangle_b$ | 0.42 $\frac{\pi}{2}$ | 96 | 1.20 |
| $|51\rangle_a$ | $|6\rangle_a$ | $|4\rangle_b$ | 1 $\frac{\pi}{2}$ | 5 | 1.54 |
| $|52\rangle_a$ | $|7\rangle_a$ | $|3\rangle_b$ | 0.48 $\frac{\pi}{2}$ | 3 | 1.56 |
| $|53\rangle_a$ | $|8\rangle_a$ | $|2\rangle_b$ | 0.45 $\frac{\pi}{2}$ | 24 | 1.40 |
| $|54\rangle_a$ | $|9\rangle_a$ | $|1\rangle_b$ | 0.64 $\frac{\pi}{2}$ | 0.04 | 1.58 |
| $|55\rangle_a$ | $|10\rangle_a$ | $|0\rangle_b$ | $\frac{\pi}{2}$ | 0.68 | 1.57 |
| $|60\rangle_a$ | $|6\rangle_a$ | $|6\rangle_b$ | 1.32 $2.8 \frac{\pi}{2}$ | 88 | 1.20 |
| $|61\rangle_a$ | $|7\rangle_a$ | $|5\rangle_b$ | 1.32 $3.16 \frac{\pi}{2}$ | 4 | 1.54 |
| $|62\rangle_a$ | $|8\rangle_a$ | $|4\rangle_b$ | 0.38 $\frac{\pi}{2}$ | 2 | 1.56 |
| $|63\rangle_a$ | $|9\rangle_a$ | $|3\rangle_b$ | 0.42 $\frac{\pi}{2}$ | 0.07 | 1.58 |
| $|64\rangle_a$ | $|10\rangle_a$ | $|2\rangle_b$ | 2.68 $\frac{\pi}{2}$ | 0.23 | 1.58 |
| $|65\rangle_a$ | $|11\rangle_a$ | $|1\rangle_b$ | 2.56 $\frac{\pi}{2}$ | 0.066 | 1.58 |
| $|66\rangle_a$ | $|12\rangle_a$ | $|0\rangle_b$ | 1.56 $\frac{\pi}{2}$ | 0.2 | 1.58 |

Fig. 8. Minimum error probability ($P_e$) vs $n$ for $M=3$ and for all forms of input states. The effective spin $|j, j-1\rangle_z$ shows the best performance.

The problem of discriminating between more than three phases grows exponentially in size, and low error probabilities appear elusive for FSI with $M>3$.

### III. QUANTUM READING

We now examine an application of binary phase discrimination, which is the reading of digital classical memory using quantum light [20][21].

We already know that FSI enables binary discrimination of smaller phase shifts than any other method, and also yields zero error probability at the few photon level, beating the best classical protocols. Here, we compare the channel capacity and photon information efficiency (PIE) of quantum optical reading to that of classical reading and show that,

also that the error probability tends to be higher than for other cases, even though it does decrease with increasing $j$. Mutual information follows the same trend, being lowest for $m=0$ and increasing with $j$. In fact, the $m=j-1$ input offers minimal error probability and maximal mutual information for each given value of $j$, as plotted in Fig.[5] This
whereas standard optical drives using a laser probe and direct detection are limited to a PIE of 0.5 bit of information per transmitted photon, FSI can reach 1 bit of information per transmitted photon for binary phase discrimination, and \( \log_2(3) \approx 1.6 \) bit of information per 3-phase pixel in ternary discrimination using as few as two photons.

The classical capacity of optical reading is the amount of bits of information that can be reliably encoded and read per pixel and is equivalent to the maximum attainable mutual information between the applied phase shifts \( \hat{\theta} \) and the measured phase shifts \( \theta \) for each pixel (see Appendix B),

\[
C(n_s) = \max \, I(\theta; \hat{\theta}),
\]

where \( n_s \) is the average number of signal photons in the reading probe, here the interferometer arm that contains the phase shift. It is straightforward to show that

\[
n_s = \langle n_a \mid \langle n_b \mid U_{BS}^\dagger a^\dagger a U_{BS} \mid n_a \rangle \mid n_b \rangle = j.
\]

The PIE is then the number of bits read per signal photons:

\[
PIE = \frac{C(n_s)}{n_s}.
\]

Here we consider a binary phase shift keyed (BPSK) phase encoding (different from the AM scheme used in optical disks) and apply binary phase discrimination by FSI to optical reading.

The mutual information (MI) for FSI with \( n_s = 1, 2 \) photons is plotted in Fig. 9. From these results, we can determine the phases that maximize the MI. Consistent with error probability studies in the previous sections, we find that the \( m = 0 \) states yield maximum MI for the smallest minimum phase. In the following, we use the MI at these optimum phases to calculate the channel capacity and PIE.

The classical capacity of a quantum channel is limited by the Holevo bound, which imposes an upper limit on the continued reliable rate of reading classical information from a quantum channel \cite{22, 23}. For the lossless phase only encoding the Holevo capacity can be written as \cite{24}

\[
C(n_s) = (1 + n_s) \log_2(1 + n_s) - n_s \log_2(n_s).
\]

In order to reach the Holevo capacity bound, one needs to use an optimum probe state as well as an optimum receiver design. However, no feasible experimental scheme is known to achieve this limit. As for binary phase discrimination, we compare FSI performance with coherent-state encoding with homodyne, heterodyne, or Dolinar receivers. For these binary channels the capacity is given by

\[
C(n_s) = 1 + P_c \log_2(P_c) + (1 - P_c) \log_2(1 - P_c).
\]

where \( P_c \) for homodyne and Dolinar is given by Eqs. (14-15). In Fig. 10 we plot the PIE versus \( n_s \) for these different approaches. As can be seen from the figure, the PIE of FSI slightly outperforms the Dolinar receiver for \( n_s = 1 \).

In Fig. 11 we plot the PIE versus the channel capacity for the same BPSK receivers and also add the coherent-state on-off encoding, which is the currently used (AM) technology for reading optical disks. Again, FSI outperforms all other receivers.

Finally, we examine the use of ternary phase shifted key (TPSK) modulation, where information is encoded as three distinct phases shifts \((0, \theta_1, \theta_2)\) in each memory pixel for optical reading, which could result in denser encoding. From our study of ternary phase discrimination in this paper, we already know the following: while an \( m = 0 \) input will resolve smaller phases, it also does that at higher probability of error than other, “less quantum” inputs such as \( m = j \). The latter, however, gives lower error probability for larger phase shifts, with \( m = j - 1 \) being optimum. The results of Table II also show that this is true for MI, which is higher, near the maximum of \( \log_2(3) \approx 1.6 \) for \( m = j, j - 1 \) compared to \( m = 0 \). Based on this we can expect a PIE of 0.8 bit/photon for \( n_s = 2 \) for the \( |jj\rangle \) input, for a choice of phases such as \((0, \pi/2, \pi)\).

**IV. CONCLUSION**

We have studied Fock-state interferometry, whose implementation is now within reach with the coming of age of on-demand single-photon sources and photon-number-resolving detectors. We showed that FSI is well suited to the particular task of phase discrimination, rather than that of phase estimation. To begin with, FSI with twin inputs \((m = 0)\) can discriminate smaller phases than other methods; the phase magnitude also decreases with the photon number more rapidly than \( j^{-3/4} \) in the domain where our study was conducted (where \( 2j \) is the total photon number). For binary discrimination, FSI can reach zero error probability; it outperforms all previously known coherent-state-based approaches, including the Dolinar receiver, at low photon levels, and matches the performance of the specifically optimized phase-eigenstate interferometry. In the more exotic case of ternary phase distribution, FSI still provides a smaller phase shift discrimination (for \( m = 0 \) inputs), as well as very low error probability and near maximal mutual information (for \( m = j - 1, j \)). We finally examined the application of phase discrimination to phase-shift-keyed optical reading and confirmed that the photon information efficiency of FSI exceeds all previously known experimentally feasible approaches, as well as the industry standard on-off AM reading.
Appendix A: The Schwinger Representation

The Schwinger representation \[13\] introduces a mathematical description of passive lossless four-port optical devices based on rotations in an abstract 3D space. The application of the Schwinger representation to the analysis of optical interferometers was first demonstrated by Yurke et al. \[4\], after Wódkiewicz and Eberly presented the relevance to optics of the SU(2) and SU(1,1) groups \[25\].

Any linear passive lossless optical device with two input and two output ports can be described by a $2 \times 2$ SU(2) matrix

$$ U = \begin{pmatrix} \cos \frac{\beta}{2} e^{i(\alpha+\gamma)/2} & \sin \frac{\beta}{2} e^{i(\alpha-\gamma)/2} \\ -\sin \frac{\beta}{2} e^{-i(\alpha+\gamma)/2} & \cos \frac{\beta}{2} e^{-i(\alpha+\gamma)/2} \end{pmatrix} \quad (A1) $$

ACKNOWLEDGMENTS

This work was funded by Raytheon-BBN, under the U.S. Defense Advanced Research Projects Agency “Information in a Photon” program, and by U.S. National Science Foundation grants PHY-0855632, PHY-1206029, and PHY-1708023.

FIG. 9. Mutual information of optical reading. The binary information is encoded in optical phase shifts $(0, \theta)$; a), $n_s = j = 1$; b), $n_s = j = 2$.

FIG. 10. Photon information efficiency versus $n_s$ for various interferometry schemes.

FIG. 11. Photon information efficiency (bits per photon) vs the encoding efficiency (bits encoded per pixel) for various input states and receivers.
where $\alpha$, $\beta$ and $\gamma$ are the Euler angles. $U$ operates on the two-dimensional vector $(a, b)^T$, whose components are the annihilation operators for the two input fields at each port of the system.

The homomorphism from SU(2) to the rotation group in three dimensions, SO(3), allows us to visualize the action of two-mode optical devices, such as beam splitters and phase shifters, as rotations in 3D space. The general rotation in Eqs. (A1) is mathematically equivalent to the rotation of the following tridimensional vector $\vec{J}$ in 3D space:

$$ J = \begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} a^\dagger b + b^\dagger a \\ -i(a^\dagger b - b^\dagger a) \\ a^\dagger a - b^\dagger b \end{pmatrix}. $$ \hspace{1cm} (A2)

Components $J_x$, $J_y$ and $J_z$ follow the canonical commutation relations for quantum angular momentum operators

$$ [J_k, J_l] = i\hbar \varepsilon_{klm} J_m $$ \hspace{1cm} (A3)

where $k, l, m \in \{x, y, z\}$, and $\varepsilon_{klm}$ is the Levi-Civita symbol. So $\vec{J}$ can be deemed a quantum angular momentum, or effective spin.

The magnitude of the angular momentum $J^2$ is

$$ J^2 = J_x^2 + J_y^2 + J_z^2 = \frac{a^\dagger a + b^\dagger b}{2} \left( \frac{a^\dagger a + b^\dagger b}{2} + 1 \right) $$ \hspace{1cm} (A4)

$$ J^2 = \frac{N}{2} \left( \frac{N}{2} + 1 \right) $$ \hspace{1cm} (A5)

where

$$ N = N_a + N_b = a^\dagger a + b^\dagger b $$ \hspace{1cm} (A6)

is the total photon number operator.

Fock states $|n_a \rangle |n_b \rangle$ are therefore also eigenstates of $J^2$ and $J_z$,

$$ |j \mu \rangle_z = |n_a \rangle_a |n_b \rangle_b $$ \hspace{1cm} (A7)

with respective eigenvalues $j(j + 1)$ and $\mu$ given by the total photon number and the photon number difference

$$ j = \frac{n_a + n_b}{2}, $$ \hspace{1cm} (A8)

$$ \mu = \frac{n_a - n_b}{2} $$ \hspace{1cm} (A9)

As an example, input state of the interferometer with 2 photons in mode $a$ and vacuum in mode $b$ is identical to $|2j \rangle_a |0 \rangle_b = |jj \rangle_z$ and the total Fock state input which is required for Holland-Burnett interferometry is $|j \rangle_a |j \rangle_b = |jj \rangle_z$.

A unitary operation on the quantum fields $a$ and $b$ can be viewed as the SO(3) rotation of the corresponding spin $\vec{J}$, Eq. (A2). Any rotation of spin $\vec{J}$ can be described with the 3 Euler rotations:

$$ \vec{J}^{\text{out}} = e^{i\alpha J_x} e^{i\beta J_y} e^{i\gamma J_z} \vec{J}^{\text{in}} e^{-i\gamma J_z} e^{-i\beta J_y} e^{-i\alpha J_x} $$ \hspace{1cm} (A10)

$$ |\psi\rangle_{\text{out}} = e^{i\alpha J_x} e^{i\beta J_y} e^{i\gamma J_z} |\psi\rangle_{\text{in}} $$ \hspace{1cm} (A11)

respectively in the Heisenberg and Schrodinger pictures. In the Schwinger representation this arbitrary tridimensional rotation of the effective spin $\vec{J}$ is equivalent to the Euler angle parametrization of the SU(2) rotation of the two modes $a$ and $b$ basis, Eq. (A1). The SO(3) Euler matrix is

$$ \begin{pmatrix} c_\alpha c_\beta c_\gamma - s_\alpha s_\beta s_\gamma & -c_\gamma s_\alpha - c_\alpha c_\beta s_\gamma & c_\alpha s_\beta \\ c_\alpha s_\beta c_\gamma + c_\beta s_\alpha & c_\alpha c_\gamma - c_\beta s_\alpha s_\gamma & -s_\alpha s_\beta \\ s_\gamma s_\beta & s_\gamma c_\beta & c_\gamma \end{pmatrix} $$ \hspace{1cm} (A12)

where $c_{(\alpha, \beta, \gamma)} = \cos(\alpha/\beta, \gamma)$, $s_{(\alpha/\beta, \gamma)} = \sin(\alpha/\beta, \gamma)$.

The SO(3) matrix for the beam splitter with Fresnel coefficients $\rho = \cos(\phi/2)$ and $\tau = \sin(\phi/2)$ is

$$ \begin{pmatrix} j_{\text{out}}^x \\ j_{\text{out}}^y \\ j_{\text{out}}^z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} j_{\text{in}}^x \\ j_{\text{in}}^y \\ j_{\text{in}}^z \end{pmatrix} $$ \hspace{1cm} (A13)

The effect of the beam splitter on modes $a$ and $b$ is a rotation of effective spin $\vec{J}$ by $(-\phi)$ around the $x$ axis. The special case of the 50/50 beam splitter with $\rho = \tau = \frac{1}{\sqrt{2}}$ and $\phi = \pi/2$ corresponds to a $(-\pi/2)$ rotation around the $x$ axis:

$$ \vec{J}^{\text{out}} = e^{i\pi/2 J_x} \vec{J}^{\text{in}} e^{-i\pi/2 J_x} $$ \hspace{1cm} (A14)

$$ |\psi\rangle_{\text{out}} = e^{i\pi/2 J_x} |\psi\rangle_{\text{in}}. $$ \hspace{1cm} (A15)

The SO(3) matrix for the phase shift $\theta$ between the arms of an interferometer is

$$ \begin{pmatrix} j_{\text{out}}^x \\ j_{\text{out}}^y \\ j_{\text{out}}^z \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} j_{\text{in}}^x \\ j_{\text{in}}^y \\ j_{\text{in}}^z \end{pmatrix} $$ \hspace{1cm} (A16)

which is a rotation of effective spin around $z$ by $\theta$.

$$ \vec{J}^{\text{out}} = e^{i\theta J_z} \vec{J}^{\text{in}} e^{-i\theta J_z} $$ \hspace{1cm} (A17)

$$ |\psi\rangle_{\text{out}} = e^{i\theta J_z} |\psi\rangle_{\text{in}}. $$ \hspace{1cm} (A18)

The Mach-Zehnder Interferometer (MZI) consists of two 50/50 beam splitters, and a phase shifter. So, the effect of MZI is equivalent to a $(-\pi/2)$ rotation around $x$ axis, a $\theta$ rotation around $z$, and another $\pi/2$ rotation around $x$, which yields a $\theta$ rotation around $y$

$$ \vec{J}^{\text{out}} = e^{i\theta J_y} \vec{J}^{\text{in}} e^{-i\theta J_y} $$ \hspace{1cm} (A19)

$$ |\psi\rangle_{\text{out}} = e^{i\theta J_y} |\psi\rangle_{\text{in}}. $$ \hspace{1cm} (A20)

So the effect of MZI is equivalent to a single rotation of effective spin by $\theta$ around the $y$ axis. We are interested on the effect of MZI on Fock states, the
eigenstates of effective spin $\vec{J}_e$. The probability function $P(\mu', \mu | \theta, j)$ for the input spin $| j, \mu' \rangle$ to be measured after the interferometer as $| j, \mu \rangle$ for fixed $\theta$ and $J$ (the total photon number) can be described as a rotation matrix, which is a square matrix of dimension $2j + 1$ with general element

$$
P(\mu', \mu | \theta, j) = | \langle j, \mu' | \psi_{\text{out}} \rangle |^2 \quad (A21)$$

$$
= | \langle j, \mu' | e^{\theta J_y} | j \mu \rangle_z |^2 \quad (A22)$$

$$
= d_{\mu', \mu}^j(\theta)^2 \quad (A23)
$$

These rotation matrix elements can be expressed in terms of Jacobi polynomials

$$
d_{\mu', \mu}^j(\theta) = \left[ \frac{(j + \mu)! (j - \mu)!}{(j + \mu')! (j - \mu')!} \right]^{1/2} \left( \sin \frac{\beta}{2} \right)^{\mu - \mu'} \times \left( \cos \frac{\beta}{2} \right)^{\mu + \mu'} \times P_{j - \mu}^{(\mu - \mu', \mu + \mu')} (\cos \beta)
$$

(A24)

### Appendix B: Information theory

In this section we briefly recall some basic concepts of information theory. We will use these concepts to study the behavior of FSI, then compare it with other schemes.

#### 1. Entropy

For any probability distribution, we recall the definition of Shannon entropy. Let $X$ be a discrete random variable with the probability function, $P(x)$, $x \in X$. The Shannon entropy of the random variable $X$ is

$$
H(X) = - \sum_{x \in X} P(x) \log_2 P(x) \quad (B1)
$$

and the conditional entropy of random variables $X,Y$, is the expected value of the entropies of the conditional distributions, averaged over the conditioning random variable,

$$
H(Y|X) = \sum_{x \in X} P(x) H(Y|X=x) \quad (B2)
$$

$$
= - \sum_{x \in X} P(x) \sum_{y \in Y} P(y|x) \log_2 P(y|x)
$$

(B3)

where $P(y|x)$ is the conditional probability of measuring $y$ given that $x$ occurred.

#### 2. Mutual information

The mutual information (MI) is a measure of the amount of information that one random variable $X$ contains about another random variable $Y$, is equivalent to the reduction in the uncertainty of one random variable due to the knowledge of the other. For random variables $X$ and $Y$ with probability functions $P(x), x \in X$ and $P(y), y \in Y$ and the conditional entropies $H(X|Y)$ and $H(Y|X)$, the mutual information can be written as:

$$
I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)
$$

(B4)

One way to study the general problem of encoding the information in $M$ optical phases, or the more specified problem of optical reading is to look at the mutual information between the applied optical phases, $\theta$ and the measured phases $\hat{\theta}$. Ideally, $\hat{\theta}$ should contain all the information about $\theta$ and the mutual information $I(\theta; \hat{\theta})$ should be equal to the amount of information in random variable $\theta$, $H(\theta)$.

Let’s assume information is encoded in $M$ optical phases with equal a priori probabilities so,

$$
\theta = \{ \theta_1, \ldots, \theta_i, \ldots, \theta_M \} \quad (B5)
$$

$$
P(\theta) = \left\{ \frac{1}{M}, \ldots, \frac{1}{M}, \ldots, \frac{1}{M} \right\} \quad (B6)
$$

$$
H(\theta) = \log_2 M \quad (B7)
$$

Then, one seeks to estimate the phases $\hat{\theta}$

$$
\hat{\theta} = \{ \hat{\theta}_1, \ldots, \hat{\theta}_i, \ldots, \hat{\theta}_M \} \quad (B8)
$$

$$
P(\hat{\theta}) = \left\{ P(\hat{\theta}_1), \ldots, P(\hat{\theta}_i), \ldots, P(\hat{\theta}_M) \right\} \quad (B9)
$$

Note that $\hat{\theta}$ is not necessarily the same random variable as $\theta$ but their similarity and the overlap in their information contents, defined as the mutual information $I(\theta; \hat{\theta})$, is a good measure of success in the phase encoding problem. $I(\theta; \hat{\theta})$ can be calculated as

$$
I(\theta; \hat{\theta}) = H(\theta) - H(\theta | \hat{\theta}) \quad (B10)
$$

$$
= H(\theta) - \sum_j P(\hat{\theta}_j) \sum_i P(\theta_i | \hat{\theta}_j) \log_2 P(\theta_i | \hat{\theta}_j)
$$

(B11)

$P(\hat{\theta}_j | \theta_i)$ is the probability that one can directly extract from experiment so, we employ the Bayes theorem,

$$
P(\theta_i | \hat{\theta}_j) = \frac{P(\hat{\theta}_j | \theta_i) P(\theta_i)}{P(\hat{\theta}_j)} \quad (B12)
$$

$$
P(\hat{\theta}_j) = \sum_i P(\hat{\theta}_j | \theta_i) P(\theta_i) \quad (B13)
$$

and substitute $P(\theta_i | \hat{\theta}_j)$ with $P(\hat{\theta}_j | \theta_i)$ in $I(\theta; \hat{\theta})$:

$$
I(\theta; \hat{\theta})
$$

$$
= H(\theta) - \sum_{i,j} P(\theta_j) \frac{P(\hat{\theta}_j | \theta_i) P(\theta_i)}{P(\theta_i)} \log_2 \frac{P(\hat{\theta}_j | \theta_i) P(\theta_i)}{P(\theta_i)}
$$

(B14)
This gives
\[
I(\theta; \hat{\theta}) = \sum_i P(\theta_i) \log_2 P(\theta_i)
\]
\[
- \sum_{i,j} P(\theta_i) P(\hat{\theta}_j | \theta_i) \log_2 \frac{P(\hat{\theta}_j | \theta_i)}{\sum_k P(\hat{\theta}_j | \theta_k)}
\]
(B15)
\[
= \log_2 M - \sum_{i,j} \frac{P(\hat{\theta}_j | \theta_i)}{M} \log_2 \frac{P(\hat{\theta}_j | \theta_i)}{\sum_k P(\hat{\theta}_j | \theta_k)}
\]
(B16)

Thus \(I(\theta; \hat{\theta})_{\text{max}} = \log_2 M\), when \(P(\hat{\theta}_j | \theta_i)_{i=j} = 1\) and \(P(\hat{\theta}_j | \theta_i)_{i \neq j} = 0\). \(P(\hat{\theta}_j | \theta_i)_{i \neq j}\) is the probability of having phase \(\theta_i\) and estimating the wrong phase \(\theta_j\) which results in error.

\[1\] C. M. Caves, Quantum-mechanical radiation-pressure fluctuations in an interferometer, Phys. Rev. Lett. 45, 75 (1980).
\[2\] J.-M. Lévy-Leblond and F. Balibar, Quantics: Rudiments of Quantum Physics (North-Holland, 1990).
\[3\] C. M. Caves, Quantum-mechanical noise in an interferometer, Phys. Rev. D 23, 1693 (1981).
\[4\] B. Yurke, S. L. McCall, and J. R. Klauder, SU(2) groups in quantum-optics applications, J. Opt. Soc. Am. B 2, 458 (1985).
\[5\] M. J. Holland and K. Burnett, Interferometric detection of optical phase shifts at the Heisenberg limit, Phys. Rev. Lett. 71, 1355 (1993).
\[6\] T. Kim, O. Pfister, M. J. Holland, J. Noh, and J. L. Hall, Influence of decorrelation on Heisenberg-limited interferometry using quantum correlated photons, Phys. Rev. A 57, 4004 (1998).
\[7\] A. Luis and L. Sánchez-Soto, Quantum phase difference, phase measurements and Stokes operators, Prog. Opt. 41, 421 (2000).
\[8\] C. W. Helstrom, Quantum Detection and Estimation Theory (Mathematics in Science and Engineering, 123,(Academic Press, New York), 1976).
\[9\] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, Phys. Rev. Lett. 96, 010401 (2006).
\[10\] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, Nat. Phys. 7, 406 (2011).
\[11\] G. S. Thekka, M. E. Mycroft, B. A. Bell, C. G. Wade, A. Eckstein, D. S. Phillips, R. B. Patel, A. Buraczewski, A. E. Lita, T. Gerrits, S. W. Nam, M. Stobińska, A. I. Lvovsky, and I. A. Walmsley, Quantum-enhanced interferometry with large heralded photon-number states, npj Quantum Information 6, 89 (2020).
\[12\] M. Perarnau-Llobet, A. González-Tudela, and J. I. Cirac, Multimode fock states with large photon number: effective descriptions and applications in quantum metrology, Quantum Science and Technology 5, 025003 (2020).
\[13\] J. Schwinger, On angular momentum, U.S. Atomic Energy Commission Report. No. NYO–3071 (1952).