Linear optical quantum metrology with single photons: Experimental errors, resource counting, and quantum Cramér-Rao bounds

Jonathan P. Olson,[1,*] Keith R. Motes,[2,†] Patrick M. Birchall,[3] Nick M. Studer,[4] Margarite LaBorde,[4] Todd Moulder,[4] Peter P. Rohde,[5,4,‡] and Jonathan P. Dowling[4]

[1]*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA*
[2]*Department of Physics and Astronomy, Macquarie University, Sydney, New South Wales 2113, Australia*
[3]*Centre for Quantum Photonics, University of Bristol, Bristol BS8 1UB, United Kingdom*
[4]*Hearne Institute for Theoretical Physics and Department of Physics & Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803, USA*
[5]*Centre for Quantum Computation and Intelligent Systems, Faculty of Engineering & Information Technology, University of Technology Sydney, New South Wales 2007, Australia*

Quantum number-path entanglement is a resource for supersensitive quantum metrology and in particular provides for sub-shot-noise or even Heisenberg-limited sensitivity. However, such number-path entanglement is thought to have been resource intensive to create in the first place, typically requiring either very strong nonlinearities or nondeterministic preparation schemes with feedforward, which are difficult to implement. Recently [K. R. Motes *et al.*, Phys. Rev. Lett. **114**, 170802 (2015)], it was shown that number-path entanglement from a BosonSampling inspired interferometer can be used to beat the shot-noise limit. In this paper we compare and contrast different interferometric schemes, discuss resource counting, calculate exact quantum Cramér-Rao bounds, and study details of experimental errors.

DOI: 10.1103/PhysRevA.96.013810

I. INTRODUCTION

A substantial fraction of the efforts of theoretical computer science and physics is now invested in the discovery of postclassical devices to demonstrate quantum supremacy. Much still remains unknown about the fundamental limits and complexity of quantum computing. One well-known example of a device, which exhibits a quantum advantage over its classical counterpart, came about as a result of the discovery of the Hong-Ou-Mandel effect, enabling interferometers to estimate unknown variables with improved sensitivity [1]. It was subsequently shown that bosonic NOON states could achieve asymptotically better sensitivities than a comparable device using only classical techniques, with an increasing number of probe photons N [2]. Like universal quantum computers, however, at present the experimental overhead for producing NOON states is prohibitive, making practical implementation infeasible [3,4]. Further, although much is now known about metrology using two-mode interferometers, far less is known about larger multimode networks [5].

Recently, the advent of postclassical devices, such as BosonSampling [6,7], has drawn new interest to the capabilities of potentially more practical passive linear optical networks. These networks generate complicated number-mode entanglement across an exponentially large state space. If such an optical network is fed with uncorrelated single photons, the output probabilities in the multimode Fock basis are given by complex matrix permanents, known to be #P-hard to compute exactly and strongly believed to be computationally intractable quantities to estimate accurately [8,9]. Boson-

Sampling has already attracted much experimental interest as a simple approach for performing a truly postclassical computation [10–14].

Recently, the sensitivity of a passive linear optical setup, consisting of single photons fed into a specific multimode interferometer and photodetection at the output, was investigated [15,16]. In this paper we study a much larger class of devices and further show how the sensitivity can be maximized in this scenario given realistic constraints on the unknown phase. We show that the device achieving this optimality is not only more sensitive than the one proposed in Ref. [15], but also far easier to construct. We achieve this by using the Fisher information formalism, which provides insight into the role of each different component of the interferometer and aids an explicit calculation of the phase sensitivity for the optimal network, a result that was only postulated in Ref. [15]. Additionally, we provide an analytic calculation of the phase sensitivity for the optimal network from matrix permanents, which is an improvement on the conjectured result of Ref. [15]. Our main result shows that for $n < 7$ photons we achieve sub-shot-noise limited sensitivity with a passive multimode linear optical device with $O(n)$ optical elements. We believe that this work is experimentally achievable with the current temporal infrastructure from Ref. [14].

Although previous models for supersensitive quantum devices have shown better theoretical scaling in the limit of larger average photon number, many of these schemes admit a more pessimistic scalability from an engineering perspective. These devices generally are not robust enough under noise models, require nonlinear components, have a large overhead in state preparation, or employ measurement schemes that are difficult to implement [17,18]. For instance, a Mach-Zehnder interferometer (MZI) with a two-mode squeezed vacuum input and parity detection performs extremely well in the noiseless regime, but degrades quickly under dephasing and loss and

_____

*olson.jonathanp@gmail.com
†motesk@gmail.com
‡dr.rohde@gmail.com; http://www.peterrohde.org

013810-1

requires high-efficiency number-resolving detectors whose ranges must scale with the average photon number [19].

This paper is organized as follows. In Sec. II we describe our generalized architecture and discuss quantum Fisher information and quantum Cramér-Rao bounds. In Sec. III we investigate the choice of measurement interferometer in relation to the state preparation interferometer. In this section we also discuss various phase strategies for $\hat{\Phi}$ and find the optimal phase strategy. In Sec. IV we describe how different choices of $\hat{V}_1$ and $\hat{V}_2$ affect device sensitivity and find their optimal structure. In Sec. V we investigate how loss and dephasing errors affect the device performance.

II. ARCHITECTURE AND QUANTUM CRAMÉR-RAO BOUNDS

To begin, let us consider the quantum Fourier transform interferometer (QuFTI) architecture, as shown in Fig. 1, invented by Motes *et al.* A QuFTI consists of three particular components: an input state $|\psi_{\mathrm{in}}\rangle = |1\rangle^{\otimes n}$ of n single photons; an n-mode interferometer with a transfer matrix V_1 that performs the operation $\hat{V}_1$, followed by a generalized linear phase evolution $\hat{\Phi}$ that encodes the unknown phase φ, and a second interference with a transfer matrix V_2 (enacting the transformation $\hat{V}_2$); and coincidence photodetection at the output. We will let the phase-evolved state be defined as $\hat{\Phi}\hat{V}_1|\psi_{\mathrm{in}}\rangle \equiv |\psi_\varphi\rangle$, where $\hat{\Phi} = \exp(i\sum_j \hat{n}_j f_j \varphi)$, with $\hat{n}_i$ being the number operator for mode i. We define the measurement to be a second evolution with a transfer matrix V_2 followed by an array of on-off (non-number-resolving) photodetectors (see Fig. 1) [15]. The most compelling aspect of the Motes *et al.* architecture is that it is composed of single-photon inputs, passive linear optics, and on-off detection, all of which can be implemented on an integrated photonic chip [20]. These simplifications forego many of the technical challenges required to generate NOON states [4] and are experimentally scalable, but only provide sub-shot-noise sensitivity for a small number of modes.
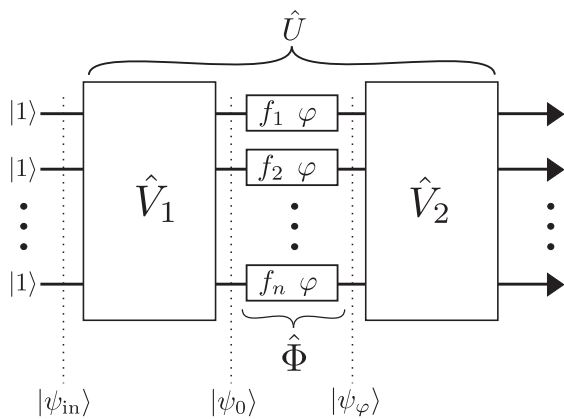

FIG. 1. Generalized architecture for the quantum Fourier transform interferometer. We consider optimizations over $V_1, V_2 \in \mathrm{SU}(n)$ and phase strategies $\hat{\Phi}$, together with single-photon inputs and photodetection in each mode. The Motes *et al.* architecture can be restored when $V_2 = V_1^\dagger$, where V_1 is the n-mode quantum Fourier transform and $f_j = (j - 1)$, with j being the mode number.

The interferometers $\hat{V}_1$, $\hat{V}_2$, and $\hat{\Phi}$ can be varied without jeopardizing the scalability of the device (see Fig. 1). Hence, the following question arises: What are the optimal choices of $\hat{V}_1$, $\hat{V}_2$, and $\hat{\Phi}$ that yield maximum phase sensitivity? In this work we answer this question.

To evaluate the phase sensitivities of different architectures, we utilize the quantum Fisher information (QFI) formalism, which we will briefly summarize. Once an unknown parameter φ has been encoded onto a quantum state ρ_φ, the QFI $\mathcal{F}(\rho_\varphi)$ bounds the achievable precision to which φ can be estimated, with an unbiased estimator, through the quantum Cramér-Rao bound (QCRB) [21],

$$\mathcal{F}(\rho_\varphi) \geqslant 1/\Delta^2\varphi, \tag{1}$$

where $\Delta^2\varphi$ is the variance in the estimate of φ. The QCRB can be saturated when one can measure multiple copies of ρ_φ in an optimal basis [22]. As such, the QFI is one way to quantify the information content of a quantum state related to an unknown parameter. If φ has been encoded into a pure probe state $|\psi_0\rangle$ by a unitary operator $|\psi_\varphi\rangle = e^{i\hat{H}\varphi}|\psi_0\rangle$, then the QFI is given by $\mathcal{F}(|\psi_\varphi\rangle) = 4(\langle\hat{H}^2\rangle_0 - \langle\hat{H}\rangle_0^2)$, where $\langle\cdot\rangle_0 = \langle\psi_0|\cdot|\psi_0\rangle$ [22].

III. PHASE STRATEGIES

We begin evaluating the Motes *et al.* framework by investigating the role of $\hat{V}_2$. We note that we measure $|\psi_\varphi\rangle = \hat{\Phi}\hat{V}_1|\psi_{\mathrm{in}}\rangle$ by sending it through the inverse optical network $\hat{V}_2 = \hat{V}_1^\dagger$ and detecting it using single-photon detectors is an optimal measurement strategy around $\varphi = 0$. This measurement strategy projects $|\psi_\varphi\rangle$ onto a basis containing $|\psi_{\varphi=0}\rangle$. Projecting onto this basis was shown to be an optimal measurement in Ref. [22]. To see this more directly, we can compute the probability P for the observable event $\hat{\mathcal{O}} = (|1\rangle\langle 1|)^{\otimes n}$ of detecting a photon at each output of the full interferometer $\hat{V}_2\hat{\Phi}\hat{V}_1$ to be $P = 1 - \varphi^2\mathcal{F}(|\psi_\varphi\rangle)/4 + O(\varphi^4)$ when $\hat{V}_2 = \hat{V}_1^\dagger$, as shown in Appendix A. This allows us to directly compute the phase sensitivity $\Delta\varphi$ obtained from the error propagation formula in terms of QFI,

$$\Delta\varphi = \frac{\sqrt{\langle\hat{\mathcal{O}}^2\rangle - \langle\hat{\mathcal{O}}\rangle^2}}{\left|\frac{d\langle\hat{\mathcal{O}}\rangle}{d\varphi}\right|}$$
$$= 1/\sqrt{\mathcal{F}(|\psi_\varphi\rangle)} + O(\varphi^2). \tag{2}$$

Thus, for $\varphi \simeq 0$, the error propagation formula saturates the QCRB and hence the measurement choice of $\hat{V}_2 = \hat{V}_1^\dagger$ is optimal; this is one of the primary results of this paper. Note that when $\hat{V}_2 = \hat{V}_1^\dagger$, as the precision calculated by the error propagation formula can be obtained by computing the QFI, the problem of maximizing the phase sensitivity of the device is equivalent to maximizing $\mathcal{F}(|\psi_\varphi\rangle)$.

We turn now to the choice of f_i, which is a particular choice of phase strategy as shown in Fig. 1. As has been noted before, by passing a probe state through an unknown phase shift k times, the effect of the phase shift is magnified [23]. If the total phase shift $k\theta$ applied to a mode can be measured to a given precision $\Delta^2(k\theta) = \alpha$, then the precision to which θ is measured is increased by a factor of k^2, so $\Delta^2\theta = \alpha/k^2$. This effect can be used to increase the precision of any procedure,

including those comprised entirely of classical techniques. Therefore, in order to compare different quantum strategies fairly, we impose the normalization condition

$$\sum_i f_i = 1, \tag{3}$$

i.e., we assume they all utilize the same total accumulated phase. Now we show that, given this normalization condition, the highest QFI can be achieved using a single-phase strategy, i.e., $f_j = \delta_{j,1}$. Analogous to the Motes *et al.* framework, the unknown phase is encoded by the unitary $\hat{\Phi} = \exp(i \sum_j \hat{n}_j f_j \varphi)$. It follows that

$$\frac{1}{4}\mathcal{F}(|\psi_\varphi\rangle) = \sum_{j,k} f_j f_k \, \mathrm{Cov}_0(\hat{n}_j, \hat{n}_k)$$

$$\leqslant \sum_{j,k} f_j f_k \, \mathrm{Var}_0(\hat{n}_h)$$

$$= \mathrm{Var}_0(\hat{n}_h), \tag{4}$$

where the covariance is $\mathrm{Cov}_0(A,B) \equiv \langle AB\rangle_0 - \langle A\rangle_0\langle B\rangle_0$ for commuting operators A and B and mode h is taken to be the one with the largest photon-number variance. If the phase shift was only put in mode h, then the QFI would simply be $4\,\mathrm{Var}(\hat{n}_h)$. We conclude that distributing the phase shift between two or more modes leads to a QFI, which is less than or equal to the QFI obtained if the phase shift is only in mode h. If we assume that the variance is largest in mode 1, which can be ensured by the choice of $\hat{V}_1$, the optimal phase distribution is $f_j = \delta_{j,1}$.

To illustrate the performance of differing phase strategies, we consider a range of functions representing trial strategies (Table I). For each phase strategy, we can use the result of Ref. [8] to numerically compute the phase sensitivity $\Delta\varphi$ using matrix permanents of $V_1 \Phi V_1^\dagger$. This technique is summarized in Appendix B and the result is plotted in Fig. 2. From this figure it is apparent that there is no improvement in phase sensitivity by distributing the phase throughout the modes and restricting φ to one mode is most effective. This result is consistent with the conclusions by Berry *et al.* as shown in Ref. [24].

With this in mind, we would like to now compare the architecture described in the original Motes *et al.* work to the phase strategy described above that optimizes the QuFTI. However, we have already made this comparison, since the work of Motes *et al.* possesses the linear phase strategy $f_j = j - 1$, whose normalized strategy is plotted against the optimal strategy in Fig. 2. This is contradictory to the

TABLE I. Functions representing trial phase strategies. Note that many of the strategies are not normalized to satisfy Eq. (3), but can easily be made so by dividing each by $\sum_{j=1}^n f_j$.

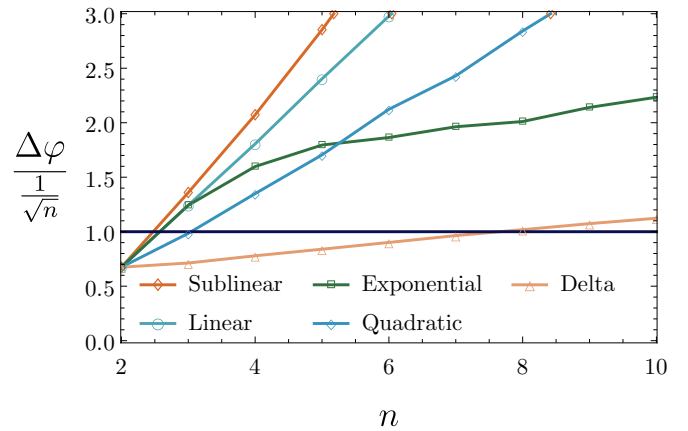| Phase strategy	Function
sublinear	$f_j^{\mathrm{sub}} = \sqrt{j}$
linear	$f_j^{\mathrm{lin}} = j - 1$
quadratic	$f_j^{\mathrm{quad}} = j^2$
exponential	$f_j^{\mathrm{exp}} = 2^j$
δ	$f_j^{\delta} = \delta_{j,1}$

FIG. 2. Phase sensitivity scaling of different phase strategies for the QuFTI suggests that widening the phase gap between modes improves the phase sensitivity. The shot-noise limit used for comparison here is defined to be $1/\sqrt{n}$, which is the best possible classical scheme for n photons and any number of modes greater than 2. The region where the sensitivity falls below one indicates supersensitivity. It is apparent that the δ-phase strategy is optimal.

preliminary results from Motes *et al.*, which showed that for all n, the phase sensitivity of Motes *et al.* beats the shot-noise limit. This is because Motes *et al.* used a different resource counting technique called ordinal resource counting (ORC). The ORC strategy found in the work of Motes *et al.* did not obey the normalization condition on the phase shifts that we have imposed here in Eq. (3). For that reason, the comparison by Motes *et al.* with the classical strategy chosen to represent the shot-noise limit was unfair. The subsequent erratum of Ref. [15] is indeed consistent with the normalized result here. Thus, when the normalization condition of Eq. (3) is imposed, the linear phase strategy used by Motes *et al.* is suboptimal.

In Fig. 3 we show the phase sensitivity of the QuFTI with the δ-phase strategy and compare it to the shot-noise limit and the Heisenberg limit. We see that we do better than the shot-noise limit for $n \leqslant 6$ photons, which is well above what
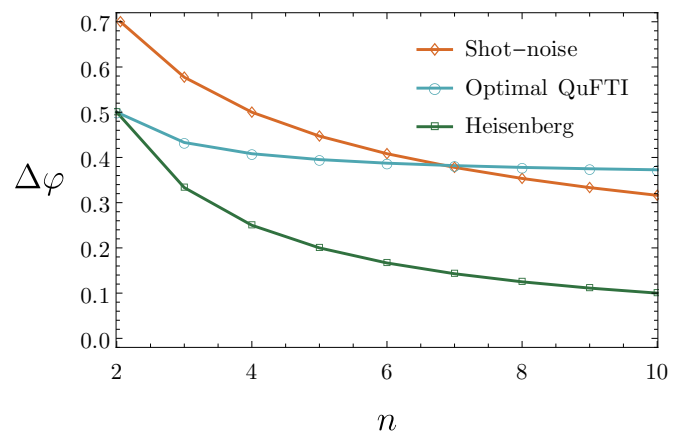

FIG. 3. Phase sensitivity of the optimal QuFTI, which consists of the δ-phase strategy for $\hat{\Phi}$, $\hat{V}_2 = \hat{V}_1^\dagger$, where $\hat{V}_1$ is the n-mode quantum Fourier transform, compared to the shot-noise limit and the Heisenberg limit.

is experimentally achievable today, suggesting that this type of quantum metrology might be the best route forward in the medium term.

IV. OPTIMAL UNITARY

In this section we investigate the effect that $\hat{V}_1$ has on the phase sensitivity, when using the previously found optimal components $\hat{V}_2 = \hat{V}_1^\dagger$ and $f_i = \delta_{1,i}$, the best phase strategy. In Sec. III it was shown that when the phase shift is placed in the first mode, i.e., $f_i = \delta_{1,i}$, then $\mathcal{F}(|\psi_\varphi\rangle) = 4 \operatorname{Var}_0(\hat{n}_1)$. For an initial state of single photons $|1\rangle^{\otimes n}$ fed into $\hat{V}_1$, we can explicitly compute, as shown in Appendix C, that $\mathcal{F}(|\psi_\varphi\rangle) = 8(1 - \sum_{i=1}^{n} |V_{1,i}|^4)$, where $V_{1,i}$ is the ith element in the top row of V_1. Physically, this means that the Fisher information is dependent only on the coupling between the input modes and the first output mode. This is perhaps to be expected, since only the first mode contains the phase to be interrogated. Additionally, we can compute the QFI when k photons $|k\rangle^{\otimes n}$ are fed into each mode of $\hat{V}_1$ to be

$$\mathcal{F}(|\psi_\varphi\rangle) = 4 \left(1 - \sum_{i=1}^{n} |V_{1,i}|^4 \right) k(k+1), \tag{5}$$

which is maximized for V_1 with $|V_{1,i}| = 1/\sqrt{n}$ for all elements on the top row. Note that while the QuFTI satisfies this constraint, it does not do so uniquely. To be as general as possible then, we consider any unitary with this structure to be "uniform" and any interferometer with these unitaries to be a quantum uniform multimode interferometer (QUMI). A QuFTI is hence a special case of a QUMI. Remarkably, because the phase sensitivity of the device relies only on the values of the first row of the matrix V_1, a network can attain the maximum sensitivity with only $O(n)$ beam splitters; this is a significant improvement over the Motes *et al.* architecture's QuFTI, which requires $O(n^2)$ beam splitters to implement. A simple implementation of the architecture introduced here is shown in Fig. 4, where the reflectivity amplitude of the beam splitter acting on modes 1 and k should be $1/\sqrt{k}$. Setting $|V_{1,i}| = 1/\sqrt{n}$ gives

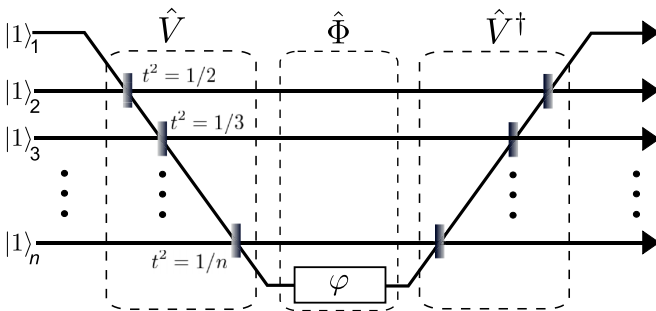$$\mathcal{F}_{\max}(|\psi_\varphi\rangle) = 4k(k+1)(1 - 1/n). \tag{6}$$

FIG. 4. Simple architecture that maximizes the phase sensitivity of our scheme. Subscripts on the kets denote mode number. The beam splitters (vertical gray slabs) should be adjusted so that V_1 is a QUMI, namely, the transmissivity amplitude t of the beam splitter acting on mode 1 and mode k should be $1/\sqrt{k}$.
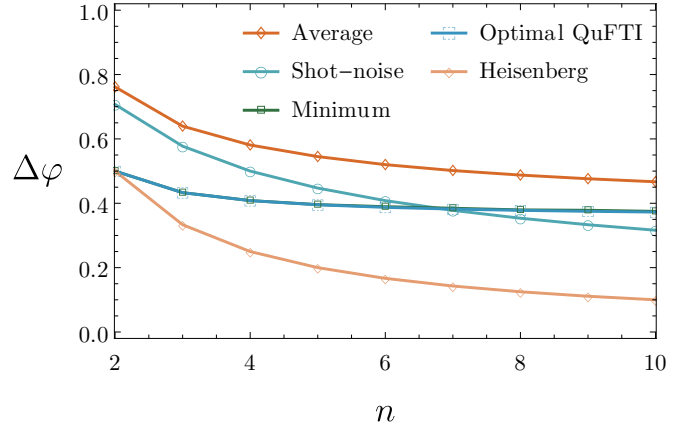
FIG. 5. The quantum Fourier transform is optimal for the δ-phase strategy. However, it is not uniquely so; we know that any uniform unitary is also optimal for this strategy.

This reduces to the result of Holland and Burnett when $n = 2$ modes [25]. Members of our team have shown that it is possible to make k-photon Fock input states from only single photons using more advanced devices, such as reliable quantum memory, which is beyond current experimental techniques [26].

For an input state of single photons the optimal precision obtainable is therefore

$$\Delta\varphi = 1/\sqrt{8(1 - 1/n)}. \tag{7}$$

One can also arrive at the same result by an explicit calculation of matrix permanents when $V = V_1 = V_2^\dagger$, which we provide in Appendix D.

To investigate the sensitivity of a device in which $\hat{V}_2 = \hat{V}_1^\dagger$ but wherein $\hat{V}_1$ is not optimal, we computed the phase sensitivity of 10 000 random unitaries in $SU(n)$ (for each n) and plotted the best phase sensitivity (i.e., minimum $\Delta\varphi$) and average phase sensitivity of this set against the phase sensitivity of the optimal QUMI (see Fig. 5). It is now clear that the best strategy is to use the δ-phase function and the QUMI. We call this overall best strategy, which has the optimal δ-phase shift, combined with the uniform weighted first row of the unitaries the optimal QUMI. Note that, although an experimental implementation of the QuFTI is not optimal for QUMIs, we use the QuFTI for analytic calculations and numerics, since the matrix itself has useful symmetries and still produces the same output statistics as any QUMI.

V. EXPERIMENTAL ERRORS

Quantum states such as single photons are notoriously difficult to manipulate. It is therefore important to consider how various errors affect the quality of metrology protocols. In particular we look at loss and dephasing. In particular, such effects generate mixed states for which the QFI is very difficult to calculate. In addition, the QFI is not practically instructive if one cannot perform an optimal measurement due to technological limitations.

For example, when loss has been applied to a quantum state, the optimal measurement strategies can require exotic

techniques such as nondemolition measurement and feedforward [18]. Therefore, calculating the QFI is not necessarily a good indicator of the performance of a practical strategy. Because of these issues, we study the performance of the architecture by directly calculating the sensitivity of specific measurement outcomes from implementable measurement techniques.

A. Loss analysis

Loss is a considerable issue to overcome in any experiment utilizing single photons. One hurdle when considering photon loss in the scheme presented here is that, if the device relies solely on bucket photon detectors, a loss event is indistinguishable from a photon collision event. Hence, for small φ, it may be the case that the loss dominates the number of perceived collision events and one is unable to obtain any useful information about φ. Employing limited photoresolution (i.e., detectors with the capability of distinguishing between one and two or more photons) can partially solve this issue, though photon losses corresponding to one of the photons in a collision event will still continue to degrade the signal. Of course, the advantage of being able to employ simple photodetectors is lost in the case that one retains the exact architecture previously presented. However, it is possible to implement pseudo-number-resolving detectors for small numbers of photons by simply coupling the output modes with one or more additional vacuum modes via beam splitters; the output of these additional modes then can use the simpler photodetector. Members of our team have also done an exhaustive calculation of loss in the context of BosonSampling in [27,28], where both spatial and temporal losses are considered.

If enough is known about the error profile beforehand, one can still extrapolate information about φ despite the noisy data. For example, if each photon has an independent and equal chance of being lost, then the sensitivity $\Delta\varphi$ degrades continuously according to the photon fidelity ℓ. Suppose $P(\varphi)$ corresponds to the probability of measuring the $|1\rangle^{\otimes n}$ outcome. If at most a single photon is lost, the probability of detecting an unambiguous collision event becomes

$$\mathrm{Pr(collision)} = (1 - P)[\ell^n + (n - 2)(1 - \ell)\ell^{n-1}]. \quad (8)$$

This can then be substituted into the error propagation formula of Eq. (2) to give the sensitivity $\Delta\varphi$ (see Fig. 6).

Recent experimental advancements have shown that linear optical networks, particularly for networks of the size considered here, can be constructed with very low loss. Networks (with even more modes and optical elements than we consider here) have been demonstrated with up to 99% efficiency [29]; additionally, single-photon detectors have achieved up to 93% efficiency [30]. Given the rapid advances in engineering optical networks, photon loss may soon be a negligible source of error for the devices considered in this paper.

B. Dephasing analysis for the optimal QuFTI

Another type of error often present in optical networks is dephasing. Here we will analyze dephasing in the optimal
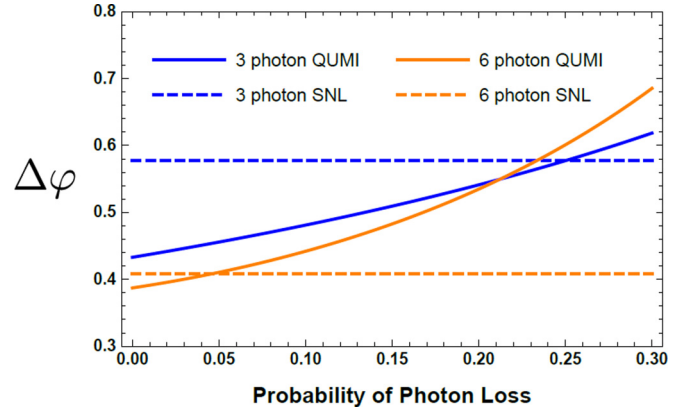

FIG. 6. Sensitivity of lossy three- and six-photon QUMIs (solid lines) compared to their respective shot-noise limits (dashed lines) at $\varphi = 0.001$. The probability on the x axis corresponds to the loss rate for each photon independently in the device.

QuFTI architecture and compare it to NOON states in a standard Mach-Zehnder interferometer.

Dephasing with a single phase shift is modeled with dephasing only occurring on the mode with the unknown phase. In the rest of the interferometer, dephasing can be made very close to zero. We insert a random phase shift χ in the single mode, which is a Gaussian random variable of zero mean $\langle\chi\rangle = 0$ but nonzero second-order moment $\langle\chi^2\rangle = \Delta\chi$,

$$e^{\pm i\varphi} \mapsto e^{\pm i(\varphi + \chi)}. \quad (9)$$

When both $\varphi, \chi \ll 1$, the approximate form of P derived in Appendix D becomes

$$P = 1 - \frac{2n - 2}{n}(\varphi + \chi)^2 + O(\varphi^4) \quad (10)$$

and correspondingly

$$\langle P \rangle \approx 1 - \frac{(2n - 2)(\varphi^2 + \Delta\chi)}{n}, \quad (11)$$
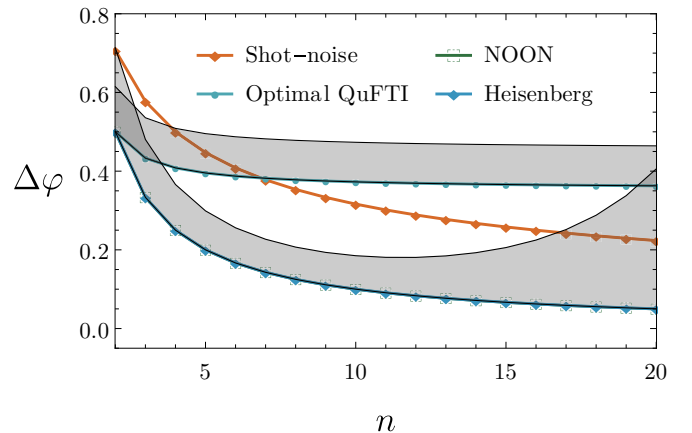

FIG. 7. Dephasing with the optimal QuFTI for $\varphi = 0.1$. This is compared to the shot-noise limit, the Heisenberg limit, and the NOON state with dephasing [31]. The shaded regions represent the dephasing regime of $0 \leqslant \Delta\chi \leqslant 0.005$ for both the optimal QuFTI and NOON states.

which is then substituted into Eq. (2) to derive the phase sensitivity. Using this result, we numerically plot the phase sensitivity with dephasing in Fig. 7 for $0 \leqslant \Delta\chi \leqslant 0.005$ and $\varphi = 0.1$. We see that, in the sub-shot-noise regime, the optimal QuFTI interferometer is comparable in dephasing to the NOON state in a MZI, another well-known metrological scheme.

It is important to note that when $\Delta\chi$ is close to or larger than the value of $\sqrt{\varphi}$, the estimator P is very poor because it is unable to differentiate between a positive or negative value of $\varphi + \chi$. Indeed, for $\varphi = 0$ and $\Delta\chi \neq 0$, the formula in Eq. (2) does not converge since $\langle P \rangle_{\varphi=0} \neq 1$. A straightforward solution is to use a known controlled phase to shift the average phase far enough away from the peak of P so that, together with the noise, the phase is predominantly positive or negative.

VI. CONCLUSION

We have considered a variety of different phase and unitary strategies for implementing a passive, single-photon input multimode metrological scheme. We have shown that the optimal architecture for n single photons is a QUMI, which equally couples each mode to a single-phase resource in one of the arms of the interferometer, followed by the inverse of the QUMI. For $n < 7$, the sensitivity of the optimal QUMI is sub-shot-noise. In the asymptotic limit of n photons, the sensitivity approaches a constant. This limit, however, assumes all n photons are used in a single experiment. The scheme can always be made to beat the shot-noise limit asymptotically by choosing an architecture with $n < 7$ modes and repeating the experiment many times.

Perhaps the most attractive feature of our proposed device is that it can be implemented with only passive linear optics, single-photon sources, and on-off detectors. In addition, it does not require any of the complicated phase arrangements from the scheme of Motes *et al.* The technology to implement the optimal architecture is essentially identical to that of BosonSampling and is achievable with experiments that have already been implemented. For a discussion on how this compares to quantum metrology using squeezed light, see Appendix E.

ACKNOWLEDGMENTS

APPENDIX A: MOTES *et al.* USE OPTIMAL MEASUREMENT

Here we link the analysis of Motes *et al.* to the quantum Fisher information formalism and point out that the measurement strategy employed is optimal. To start, we calculate the probability that all photons exit the same mode that they enter,

$$
\begin{aligned}
p(k, \ldots, k) &= |\langle k|^{\otimes m}|\hat{V}^{-1}[\mathbb{I} + i\varphi\hat{H} - \varphi^2\hat{H}^2/2 \\
&\quad - i\,O(\varphi^3)]\hat{V}|k\rangle^{\otimes k}|^2 \\
&= |1 + i\varphi\langle\hat{H}\rangle - \varphi^2\langle\hat{H}^2\rangle/2 - i\,O(\varphi^3)|^2 \\
&= 1 - \varphi^2(\langle\hat{H}\rangle^2 - \langle\hat{H}^2\rangle) + O(\varphi^4) \\
&= 1 - \varphi^2\mathcal{F}(|\psi\rangle)/4 + O(\varphi^4),
\end{aligned} \tag{A1}
$$

where we have denoted the unitary performed by the linear optical network as $\hat{V}$. Entering $p(k, \ldots, k)$ into the error propagation formulas,

$$
\Delta\varphi = \frac{\sqrt{P^2 - P}}{\left|\frac{\partial P}{\partial\varphi}\right|}, \tag{A2}
$$

and evaluating at $\varphi = 0$ gives

$$
\Delta\varphi = \frac{1}{\sqrt{\mathcal{F}(|\psi\rangle)}}. \tag{A3}
$$

Hence, this measurement basis is optimal as the quantum Cramér-Rao bound is saturated.

In fact, in a footnote in Ref. [22] it was noted that a measurement strategy, which projected onto the state $\exp(i\varphi\hat{H})|\psi\rangle$, would indeed be optimal. This is effectively the measurement that is being performed in the Motes *et al.* framework making an explicit calculation, as shown here, somewhat redundant.

APPENDIX B: CONNECTION BETWEEN $\Delta\varphi$ AND MATRIX PERMANENTS

Here we summarize how matrix permanents can be used to compute the phase sensitivity $\Delta\varphi$ of an interferometer of the form of Fig. 1. We wish to numerically compute the probability P of measuring a single photon in each mode, which is the observable $\langle\hat{\mathcal{O}}\rangle$ by which we obtain an estimate of φ. Since the input state and measurement is fixed across every strategy, we can use the result of Ref. [8] to compute $P = |\text{perm}(\hat{U})|^2 = |\text{perm}(\hat{V}\hat{\Phi}\hat{V}^\dagger)|^2$, where perm($\cdot$) refers to the matrix permanent given by the equation

$$
\text{perm}(\hat{U}) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} u_{i,\sigma(i)}, \tag{B1}
$$

with S_n the symmetric group of degree n. The phase sensitivity $\Delta\varphi$ is then found by applying the formula for standard error propagation as in Eq. (2), which can be rewritten in terms of P as

$$
\Delta\varphi = \frac{\sqrt{P - P^2}}{\left|\frac{dP}{d\varphi}\right|}. \tag{B2}
$$

APPENDIX C: OPTIMAL STATE PREPARATION

Here we calculate the QFI for the Motes *et al.* setup with k photons entering each mode and the phase shift confined to a single mode. As before, let $\hat{V} = \hat{V}_1 = \hat{V}_2^\dagger$. Putting the phase shift in the first mode, the QFI can be calculated using the Heisenberg picture. Labeling the modes in between $\hat{V}$ and $\hat{V}^\dagger$ as b_i, where $i = 1, 2, \ldots, m$, then the generator of the phase shift is $\hat{H} = \hat{b}_1^\dagger\hat{b}_1$. By utilizing $\hat{b}_i^\dagger = \sum_j V_{i,j}\hat{a}_j^\dagger$ [32] it is clear

that only the top row of V is important. Now we evaluate

$$
\langle\psi|\hat{b}_1^\dagger\hat{b}_1\hat{b}_1^\dagger\hat{b}_1|\psi\rangle = \sum_{i,j,q,l} V_{1,i}\bar{V}_{1,j}V_{1,q}\bar{V}_{1,l}\langle k|^{\otimes m}\hat{a}_i^\dagger\hat{a}_j\hat{a}_q^\dagger\hat{a}_l|k\rangle^{\otimes m}
$$

$$
= \sum_{q,l,q\neq l}|V_{1,q}|^2|V_{1,l}|^2\langle k|^{\otimes m}\hat{a}_l^\dagger\hat{a}_q\hat{a}_q^\dagger\hat{a}_l|k\rangle^{\otimes m} + \sum_{q,l}|V_{1,q}|^2|V_{1,l}|^2\langle k|^{\otimes m}\hat{a}_l^\dagger\hat{a}_l\hat{a}_q^\dagger\hat{a}_q|k\rangle^{\otimes m}
$$

$$
= \sum_{q,l,q\neq l}|V_{1,q}|^2|\bar{V}_{1,l}|^2 k(k+1) + k^2 = \sum_q\left(1-|V_{1,q}|^2\right)|\bar{V}_{1,q}|^2 k(k+1) + k^2
$$

$$
= \left(1-\sum_q|V_{1,q}|^4\right)k(k+1) + k^2 \tag{C1}
$$

and similarly

$$
\langle\psi|\hat{b}_1^\dagger\hat{b}_1|\psi\rangle = \sum_{q,l} V_{1,q}\bar{V}_{1,l}\langle k|^{\otimes m}\hat{a}_q^\dagger\hat{a}_l|k\rangle^{\otimes m}
$$

$$
= \sum_q|V_{1,q}|^2\langle k|^{\otimes m}\hat{a}_q^\dagger\hat{a}_q|k\rangle^{\otimes m}
$$

$$
= k. \tag{C2}
$$

The QFI is

$$
4\left(1-\sum_q|V_{1,q}|^4\right)k(k+1). \tag{C3}
$$

So to maximize the QFI, $\sum_q|V_{1,q}|^4$ should be minimized, which is achieved for $|V_{1,q}| = 1/\sqrt{m}$, giving a QFI of

$$
4\left(1-1/m\right)k(k+1). \tag{C4}
$$

When the number of modes equals 2 this reduces to the case studied by Holland and Burnett [25]. We note that the only part of V that played a role in this calculation was the magnitudes of the elements in the top row. Therefore, instead of a QFT circuit, a series of $m-1$ beam splitters will also be optimal.

APPENDIX D: DERIVATION OF $\Delta\varphi$ FROM MATRIX PERMANENTS

In order to derive the analytic form of $\Delta\varphi$ from Eq. (B2), we first need the matrix entries of the entire network $\hat{U} = \hat{V}\hat{\Phi}\hat{V}^\dagger$. For a single unknown phase shift φ in the first mode, $\hat{\Phi}$ has the matrix form

$$
\Phi_{j,k} = \delta_{j,k}(e^{i\varphi})^{\delta_{j,1}}. \tag{D1}
$$

Although any choice of uniform $\hat{V}$, such that $|V_{j,1}| = 1/\sqrt{n}$, should be optimal for sensitivity, for this derivation we will choose $\hat{V}$ to be the n-multimode QuFTI, which shares this property. The matrix entries of the entire network become

$$
U_{j,k} = (V\Phi V^\dagger)_{j,k} = \sum_{l,m=1}^n V_{j,l}\Phi_{l,m}V_{m,k}^\dagger = \sum_{l,m=1}^n \underbrace{\frac{1}{\sqrt{n}}\omega_n^{(j-1)(l-1)}}_{V_{j,l}}\underbrace{\delta_{l,m}e^{i\varphi\delta_{l,1}}}_{\Phi_{l,m}}\underbrace{\frac{1}{\sqrt{n}}\omega_n^{(m-1)(1-k)}}_{V_{m,k}^\dagger}
$$

$$
= \frac{1}{n}\left[e^{i\varphi} + \sum_{l=2}^n \omega_n^{(j-1)(l-1)}\omega_n^{(l-1)(1-k)}\right] = \frac{1}{n}\left[e^{i\varphi} + \sum_{l=2}^n(\omega_n^{(j-k)})^{(l-1)}\right] = \frac{1}{n}\left[e^{i\varphi} + \sum_{l=1}^{n-1}(\omega_n^{(j-k)})^l\right] \tag{D2}
$$

$$
= \begin{cases}\frac{1}{n}[e^{i\varphi} + n - 1] & j=k \\ \frac{1}{n}[e^{i\varphi} - 1] & j\neq k\end{cases} = \frac{1}{n}[e^{i\varphi} + (\delta_{j,k})n - 1]. \tag{D3}
$$

For $j=k$, it is easy to see the sum in Eq. (D2) should be $n-1$ since each term is simply $1^l = 1$. For $j\neq k$, the result follows from the fact that the sum of all nth roots of unity is zero,

$$
0 = \sum_{l=1}^n\omega_n^l. \tag{D4}
$$

The proof for the above follows directly from the geometric series and it easy to see that it extends to a sum over $\omega_n^{(j-k)}$ as well.

Now that we have the matrix entries of the network, we can compute the permanent of $\hat{U} = \hat{V}\hat{\Phi}\hat{V}^\dagger$ which is, by definition,

$$
\mathrm{perm}(U) = \sum_{\sigma\in S_n}\prod_{j=1}^n\frac{1}{n}[e^{i\varphi} + (\delta_{j,\sigma(j)})n - 1]
$$

$$
= \frac{1}{n^n}\sum_{\sigma\in S_n}\prod_{j=1}^n[e^{i\varphi} + (\delta_{j,\sigma(j)})n - 1]. \tag{D5}
$$

Suppose σ_k is some permutation with k fixed points, recalling that a *fixed point* of a permutation σ is a value $j\in\{1,..,n\}$ such that $\sigma(j) = j$ (also referred to as a partial derangement);

then the product in Eq. (D5) corresponding to σ_k is

$$\prod_{j=1}^{n}[e^{i\varphi} + (\delta_{j,\sigma_k(j)})n - 1] = [e^{i\varphi} + n - 1]^k[e^{i\varphi} - 1]^{n-k}.$$
(D6)

The sum in Eq. (D5) can now be rewritten in terms of a sum over the number of fixed points in a permutation, whose coefficient $D_{n,k}$ enumerates all permutations in S_n with k fixed points. The quantities $D_{n,k}$ are referred to as the *rencontres numbers*,

$$D_{n,k} = \frac{n!}{k!}\sum_{j=0}^{n-k}\frac{(-1)^j}{j!}.$$
(D7)

The permanent is thus

$$\text{perm}(U) = \frac{1}{n^n}\sum_{k=0}^{n}D_{n,k}[e^{i\varphi} + n - 1]^k[e^{i\varphi} - 1]^{n-k}.$$
(D8)

We are mostly interested in the behavior of $\text{perm}(U)$ for small φ, where the phase sensitivity is optimal. To simplify the remaining calculations, we focus our attention on the Taylor expansion of $F_n[\varphi] \equiv \text{perm}(U)$ around the point $\varphi = 0$, up to second order,

$$F_n[\varphi] \approx F_n[0] + F_n'[0]\varphi + \frac{1}{2}F_n''[0]\varphi^2.$$
(D9)

We can find $F_n[0]$ easily by noting that, because of the product with $[e^{i\varphi} - 1]^{n-k}$ the only nonzero term in Eq. (D8) corresponds to $k = n$,

$$F_n[0] = \frac{1}{n^n}D_{n,n}[1 + n - 1]^n = 1.$$
(D10)

Similarly, the only nonzero terms in $F_n'[0]$ must be derivatives of either $k = n$ or $k = n - 1$. Since $D_{n,n-1} = 0$, we need only concern ourselves with the derivative of the $k = n$ term. Applying the chain rule gives

$$F_n'[0] = \left[\frac{1}{n^n}D_{n,n}[e^{i\varphi} + n - 1]^n\right]'_{\varphi=0}$$

$$= \left[\frac{1}{n^n}D_{n,n}n[e^{i\varphi} + n - 1]^{n-1}ie^{i\varphi}\right]_{\varphi=0}$$
(D11)

$$= \left[\frac{1}{n^n}1 \times n[1 + n - 1]^{n-1}i\right]$$
(D12)

$$= \frac{n^n}{n^n}i = i.$$
(D13)

Evaluating $F_n''[0]$ is only marginally more difficult. The $k = n$ term can be evaluated by straightforward application of the product rule to Eq. (D11). Also, although the second derivative of the $k = n - 2$ term may be nonzero and contains a product, it is only so for the second derivative of $[e^{i\varphi} - 1]^2$; the other terms originating from the product rule are zero. Hence, $F_n''[0]$ has only three nonzero terms

$$F_n''[0] = \left[\frac{1}{n^n}D_{n,n}n[e^{i\varphi} + n - 1]^{n-1}ie^{i\varphi}\right]'_{\varphi=0} + \left[\frac{1}{n^n}D_{n,n-2}[e^{i\varphi} + n - 1]^{n-2}[e^{i\varphi} - 1]^2\right]''_{\varphi=0}$$

$$= \left[\frac{1}{n^n}D_{n,n}n(n-1)[e^{i\varphi} + n - 1]^{n-2}(ie^{i\varphi})^2\right]_{\varphi=0} + \left[\frac{1}{n^n}D_{n,n}n[e^{i\varphi} + n - 1]^{n-1}(ie^{i\varphi})^2\right]_{\varphi=0}$$

$$+ \left[\frac{1}{n^n}D_{n,n-2}2[e^{i\varphi} + n - 1]^{n-2}(ie^{i\varphi})^2\right]_{\varphi=0}$$

$$= -\left[\frac{1}{n^n}(n-1)n^{n-1}\right] - \left[\frac{1}{n^n}n^n\right] - \left[\frac{1}{n^n}2D_{n,n-2}n^{n-2}\right]$$

$$= -\left[\frac{n-1}{n} + 1 + \frac{2D_{n,n-2}}{n^2}\right] = -\left[\frac{n-1}{n} + 1 + \frac{n(n-1)}{n^2}\right]$$

$$= -\left[\frac{2n-2}{n} + 1\right] = -\frac{3n-2}{n}.$$
(D14)

Thus, Eq. (D9) becomes the simple expression

$$\text{perm}(U) \approx 1 + i\varphi - \left(\frac{3n-2}{2n}\right)\varphi^2.$$
(D15)

Recall that the probability of observing n photons, each exiting individual ports, is $P = |\text{perm}(\hat{U})|^2$. For small φ then,

$$P = \left|1 + i\varphi - \left(\frac{3n-2}{2n}\right)\varphi^2\right|^2 = \left[1 + i\varphi - \left(\frac{3n-2}{2n}\right)\varphi^2\right]\left[1 - i\varphi - \left(\frac{3n-2}{2n}\right)\varphi^2\right]$$

$$= 1 + i\varphi - i\varphi - 2\left(\frac{3n-2}{2n}\right)\varphi^2 - i^2\varphi^2 + O(\varphi^4) = 1 - \frac{2n-2}{n}\varphi^2 + O(\varphi^4).$$
(D16)

Finally, $\Delta\varphi$ becomes

$$\Delta\varphi = \frac{\sqrt{P - P^2}}{\left|\frac{\partial P}{\partial \varphi}\right|} = \frac{\sqrt{1 - \frac{2n-2}{n}\varphi^2 - 1 + \frac{4n-4}{n}\varphi^2}}{\frac{4n-4}{n}\varphi} = \frac{\sqrt{\frac{2n-2}{n}\varphi^2}}{2\frac{2n-2}{n}\varphi} = \frac{1}{2\sqrt{2}\sqrt{\frac{n-1}{n}}},$$
(D17)

which is in agreement with Eq. (7). The ratio between $\Delta\varphi$ and the shot-noise-limited phase sensitivity for n photons is then

$$\frac{\Delta\varphi}{1/\sqrt{n}} = \sqrt{\frac{n^2}{8(n-1)}}, \qquad \text{(D18)}$$

which is greater than one (i.e., gives an advantage over shot noise) for $2 \leqslant n \leqslant 6$. This is more photons than what is experimentally available today.

APPENDIX E: COMPARISON TO METROLOGY WITH SQUEEZED LIGHT

Let us compare our digital approach to metrology here to the more analog approach of coherent light mixed with a single-mode squeezed vacuum in a balanced MZI [33]. The phase uncertainty $\Delta\varphi_\xi$ for that scheme can be written as [34]

$$\Delta\varphi_\xi = [\bar{n}_\xi \exp(2r) + 2\bar{n}_\xi(\bar{n}_\xi+1)]^{-1/2}. \qquad \text{(E1)}$$

Here $\bar{n}_\xi$ is the mean number of photons in the squeezed state and r is the usual squeezing parameter [35]. We have also taken $\bar{n}_\xi = \bar{n}_\alpha$, the latter being the mean number of photons in the coherent state, which is where the scheme is optimal [36]. This expression clearly has Heisenberg scaling in the limit of infinite squeezing [or equivalently as $\bar{n}_\xi = \sinh^2(r)$ goes to infinity].

To compare with experiments we use the handy formula [35]

$$r = \tfrac{1}{2}\ln[10^{(dB/10)}], \qquad \text{(E2)}$$

which allows us to plot $\Delta\varphi_\xi$ as a function of dB, the latter of which is the typical experimental number for determining the amount of squeezing (see Fig. 8). As we can see, $\Delta\varphi_\xi$ always remains below the shot-noise limit and in fact remains below the Heisenberg limit, approaching it from below asymptotically as $\Delta\varphi_\xi \to \frac{1}{2\bar{n}_\xi}$, which we call Heisenberg scaling. It is now well known that the Heisenberg limit can be beaten with squeezed vacuum [19]. The record squeezing for optics in the laboratory is 15 dB ($\bar{n}_\xi = 7.4$), but 10 dB ($\bar{n}_\xi = 2.03$) is more typical [37].

Hence 10 dB corresponds to a mean photon number of only 2.0 in the squeezed vacuum mixed with a coherent state scheme. If we then compare this result to our scheme here in Fig. 2 for $n = 2$, our interferometer also reduces to a MZI with a twin photon input and the result is well known to be Heisenberg limited with $\Delta\varphi = 1/2 = 0.5$. For the comparable photon number in the squeezed plus coherent state MZI,
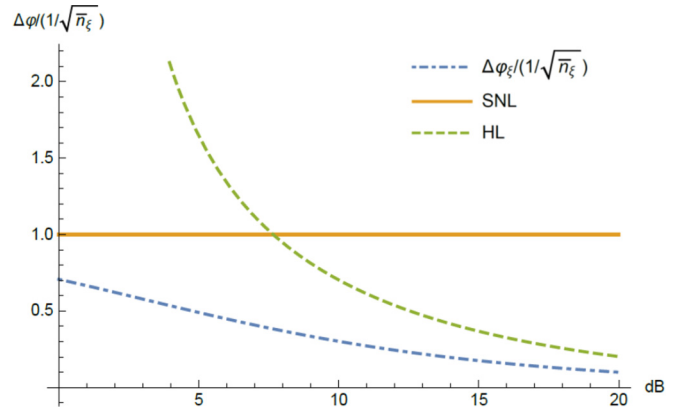

FIG. 8. Phase sensitivity of a squeezed vacuum (dot-dashed line) mixed with coherent light on a MZI.

$\Delta\varphi_\xi = 0.21$, which is slightly better but assumes that the squeezed vacuum input is in a pure state; this assumption is rarely the case. The real issue here is that while our current single-photon scheme has poor scaling at high photon number, the squeezing scheme cannot experimentally reach the high mean photon numbers required to approach the Heisenberg limit in the asymptotic regime.

Lest we be led astray, the point of our paper here is to demonstrate that the *deterministic* entanglement produced in a *linear* interferometer via the generalized Hong-Ou-Mandel effect (the same effect that generates number-path entanglement in boson sampling) can indeed provide a metrological advantage, even if small, for this particular class of interferometers. Given that groups around the world are rapidly developing optical interferometers with deterministic on-demand single-photon sources, as well as detectors, all on chip, it is reasonable to predict that systems with hundreds of single-photon inputs will soon be available. Such a chip would produce an exponentially huge amount of number-path entanglement, as has been show in the context of boson sampling [38]. If another scheme, such as multiparameter estimation or even superresolving imaging, can be found to take advantage of this entanglement, it would mean a breakthrough in quantum metrology, since only single photons, linear optics, and photon on-off detection is required. In contrast, for the squeezed vacuum approach, squeezed vacuum states with $\bar{n}_\xi = 100$ will likely never be made in the laboratory, the device that produces the squeezed vacuum in the first place requires a strong optical nonlinearity, and the detection scheme typically requires homodyning, all of which is difficult to implement on chip.

[1] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).

[2] J. P. Dowling, Contemp. Phys. **49**, 125 (2008).

[3] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, Rev. Mod. Phys. **79**, 135 (2007).

[4] H. Cable and J. P. Dowling, Phys. Rev. Lett. **99**, 163604 (2007).

[5] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Prog. Opt. **60**, 345 (2015).

[6] S. Aaronson and A. Arkhipov, in *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, San Jose, 2011* (ACM, New York, 2011), p. 333.

[7] B. T. Gard, K. R. Motes, J. P. Olson, P. P. Rohde, and J. P. Dowling, in *From Atomic to Mesoscale: The Role of Quantum Coherence in Systems of Various Complexities*, edited by S. A. Malinovaskaya and I. Novikova (World Scientific, Singapore, 2015), Chap. 8, pp. 167–192.

[8] S. Scheel, Acta Phys. Slovaca **58**, 675 (2008).

[9] L. G. Valiant, Theor. Comput. Sci. **8**, 189 (1979).

[10] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Science **339**, 794 (2013).

[11] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnola, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Nat. Photon. **7**, 545 (2013).

[12] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, Nat. Photon. **7**, 540 (2013).

[13] J. B. Spring *et al.*, Science **339**, 798 (2013).

[14] Y. He *et al.*, Phys. Rev. Lett. **118**, 190501 (2017).

[15] K. R. Motes, J. P. Olson, E. J. Rabeaux, J. P. Dowling, S. J. Olson, and P. P. Rohde, Phys. Rev. Lett. **114**, 170802 (2015).

[16] B. Opanchuk, L. Rosales-Zrate, M. D. Reid, and P. D. Drummond, arXiv:1609.05614.

[17] R. Demkowicz-Dobrzanski, J. Kolodynski, and M. Guta, Nat. Commun. **3**, 1063 (2012).

[18] R. Demkowicz-Dobrzański, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Phys. Rev. A **80**, 013825 (2009).

[19] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling, Phys. Rev. Lett. **104**, 103602 (2010).

[20] J. Carolan *et al.*, Science **349**, 711 (2015).

[21] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[22] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[23] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Nature (London) **450**, 393 (2007).

[24] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, Phys. Rev. A **80**, 052114 (2009).

[25] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).

[26] K. R. Motes, R. L. Mann, J. P. Olson, N. M. Studer, E. A. Bergeron, A. Gilchrist, J. P. Dowling, D. W. Berry, and P. P. Rohde, Phys. Rev. A **94**, 012344 (2016).

[27] K. R. Motes, J. P. Dowling, A. Gilchrist, and P. P. Rohde, Phys. Rev. A **92**, 052319 (2015).

[28] K. R. Motes, A. Gilchrist, J. P. Dowling, and P. P. Rohde, Phys. Rev. Lett. **113**, 120501 (2014).

[29] H. Wang *et al.*, Nat. Photon. **11**, 361 (2017).

[30] V. Zwiller, I. E. Zadeh, J. Los, R. Gourgues, V. Steinmetz, S. Dobrovolskiy, and S. N. Dorenbos, in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (Optical Society of America, Washington, DC, 2017), paper FF1E.1.

[31] B. Roy Bardhan, K. Jiang, and J. P. Dowling, Phys. Rev. A **88**, 023857 (2013).

[32] W. Vogel and D.-G. Welsch, *Quantum Optics* (Wiley, New York, 2006).

[33] C. Caves, Phys. Rev. D **23**, 1693 (1981).

[34] B. T. Gard, C. You, D. K. Mishra, R. Singh, H. Lee, T. R. Corbitt, and J. P. Dowling, EPJ Quantum Technol. **4**, 1 (2017).

[35] A. I. Lvovsky, in *Photonics, Volume 1, Fundamentals of Photonics and Physics*, edited by D. L. Andrews (Wiley, New York, 2015), pp. 121–164.

[36] H. F. Hofmann and T. Ono, Phys. Rev. A **76**, 031806(R) (2007).

[37] H. Vahlbruch, M. Mehmet, K. Danzmann, and R. Schnabel, Phys. Rev. Lett. **117**, 110801 (2016).

[38] B. T. Gard, J. P. Olson, R. M. Cross, M. B. Kim, H. Lee, and J. P. Dowling, Phys. Rev. A **89**, 022328 (2014).