Single-Shot Non-Gaussian Measurements for Optical Phase Estimation

M. T. DiMario and F. E. Becerra [*]

Center for Quantum Information and Control, Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131, USA

Estimation of the properties of a physical system with minimal uncertainty is a central task in quantum metrology. Optical phase estimation is at the center of many metrological tasks where the value of a physical parameter is mapped to the phase of an electromagnetic field and single-shot measurements of this phase are necessary. While there are measurements able to estimate the phase of light in a single shot with small uncertainties, demonstrations of near-optimal single-shot measurements for an unknown phase of a coherent state remain elusive. Here, we propose and demonstrate strategies for single-shot measurements for *ab initio* phase estimation of coherent states that surpass the sensitivity limit of heterodyne measurement and approach the Cramer-Rao lower bound for coherent states. These single-shot estimation strategies are based on real-time optimization of coherent displacement operations, single photon counting with photon number resolution, and fast feedback. We show that our demonstration of these optimized estimation strategies surpasses the heterodyne limit for a wide range of optical powers without correcting for detection efficiency with a moderate number of adaptive measurement steps. This is, to our knowledge, the most sensitive single-shot measurement of an unknown phase encoded in optical coherent states.

The realization of measurements for precise estimation of physical quantities is essential in physics and engineering. The fundamental limits in precision for estimating a physical parameter depend on the state used to probe the system and the process in which the parameter is encoded in the probe [1–3]. A central problem in quantum metrology is the determination of the fundamental limits on the achievable precision and their practical attainability [4–9]. Achieving such quantum measurement limits given a set of physical states using physically realizable measurements is the goal of practical estimation problems.

Optical phase estimation, where information is encoded in the phase of an electromagnetic field, is essential in tasks ranging from interferometry [5] to waveform [10] and force sensing [11,12]. Enhanced phase estimation with and without quantum states of light has been widely investigated for sensing small deviations from a known phase [13–18], and for phase estimation with repeated sampling [19,20], and with feedback measurements [14,21–25]. For these particular estimation tasks, near-optimal sensitivity for phase estimation has been approached [5,26,27].

A different and challenging problem in parameter estimation is the realization of single-shot measurements of an unknown phase. In this estimation task, a measurement is realized to estimate the phase carried by a single optical mode in a single shot [28], either with coherent fields [29,30] or quantum states of light [31]. Single-shot measurements of a completely unknown phase of a coherent state are essential for the cooling of mechanical oscillators [32,33] and the preparation of spin squeezed

states based on measurement backaction [34], as well as for high-sensitivity waveform [35] and force detection [36].

Adaptive Gaussian measurements based on homodyne detection have been extensively investigated for single-shot phase estimation with coherent states [28–30,37,38]. These schemes can in principle outperform the heterodyne measurement limit and asymptotically approach the ultimate sensitivity for optical phase estimation given by the Cramer-Rao lower bound (CRLB) [28–30]. Proof-of-principle experiments have demonstrated the potential of dyne adaptive measurements to surpass the heterodyne measurement limit for single-shot phase estimation of coherent states [39] and microwave-photon wave packets [40] after correction for experimental inefficiencies.

Here, we demonstrate optimized adaptive non-Gaussian measurements based on photon counting for single-shot phase estimation of an unknown phase of optical coherent states. These measurements use optimized coherent displacement operations, photon number resolving (PNR) detection, and conditional feedback to enable estimation strategies with high sensitivities while being robust to noise. These optimized estimation strategies allow for surpassing the heterodyne measurement limit and approaching the CRLB for coherent states with current technologies. Our experimental demonstration uses PNR detection with finite number resolution and real-time optimization of the displacement operations conditioned on the detection history as the measurement progresses. Our demonstration surpasses the heterodyne measurement limit without correcting for system inefficiencies and
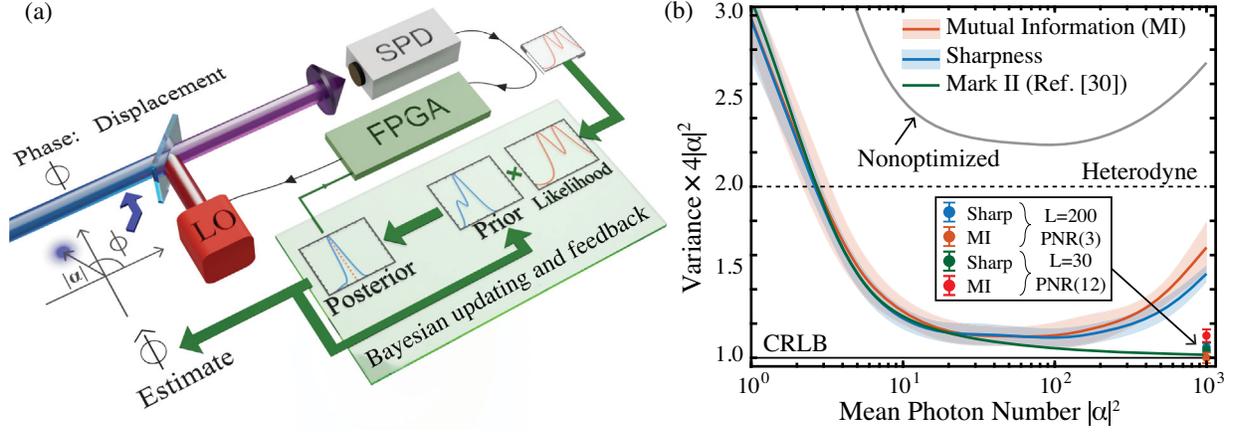
FIG. 1. Optimized non-Gaussian estimation strategy. (a) Concept of the adaptive displaced photon counting measurement for single-shot phase estimation. (b) Simulated Holevo variance multiplied by the quantum Fisher information for coherent states $(4|\alpha|^2)$. Shown are strategies maximizing the sharpness $\langle S(\beta, m)\rangle$ (blue) and the mutual information $I(\beta, m)$ (orange), and a nonoptimized strategy (gray), all with $L = 30$ adaptive steps, PNR(3), and ideal efficiency. Also shown is the CRLB for coherent states $(1/4|\alpha|^2)$, the lower bound of a heterodyne measurement $(1/2|\alpha|^2)$, and the performance of the adaptive homodyne scheme termed "Mark II" from Ref. [30]. Bold lines are the average of five Monte-Carlo simulations of 10^3 randomly distributed initial phases, and the shaded regions represent one standard deviation. The performance of both estimation strategies at $|\alpha|^2 = 10^3$ for PNR(3) with $L = 200$ and for PNR(12) with $L = 30$ are shown for reference.

approaches the CRLB for coherent states when compared to a system of the same efficiency. We believe this is the most sensitive single-shot measurement of an unknown optical phase of a coherent state to date.

Non-Gaussian phase estimation strategy.—Figure 1(a) shows the concept of the non-Gaussian strategy for single-shot phase estimation of a coherent state pulse in a single mode $|\alpha_0\rangle = |\alpha e^{i\phi_0}\rangle$ with a known mean photon number $\langle \hat{n}\rangle = |\alpha|^2$ but an unknown phase $\phi_0 \in [0, 2\pi)$. In these strategies, the displacement field is optimized in real time based on photon counting measurements within the single optical mode.

For a coherent state with unknown phase, the estimation strategy implements L adaptive measurement steps over this single mode. In the first adaptive step ($k = 1$), the input state $|\alpha_0\rangle$, with prior probability distribution for the phase $P(\phi) = 1/2\pi$, is displaced in phase space by $\hat{D}(\beta)$ using interference on a highly transmissive beam splitter to the state $\hat{D}(\beta)|\alpha_0\rangle = |\alpha_0 + \beta\rangle$. The photons in the displaced state are then detected by a single photon detector described by the operators $\hat{\Pi}_n = |n\rangle\langle n|$ for detection of n photons and $\hat{\Pi}_m = \hat{I} - \sum_{n=0}^{m-1} |n\rangle\langle n|$ for detection of m or more photons. Here m refers to the photon number resolution PNR(m) of the detector [41,42]. Given the displacement field β and detection result n, the strategy obtains the posterior probability distribution $P(\phi|n, \beta)$ for the phase through Bayes' rule: $P(\phi|n, \beta) \propto \mathcal{L}(n|\phi, \beta)P(\phi)$. Here $\mathcal{L}(n|\phi, \beta)$ is the likelihood function for detecting n photons given the displacement $\hat{D}(\beta)$:

$$\mathcal{L}(n|\phi, \beta) = \mathrm{Tr}[\hat{\Pi}_n \hat{D}(\beta)|\alpha\rangle\langle\alpha|\hat{D}^\dagger(\beta)] = |\langle n|\alpha + \beta\rangle|^2. \quad (1)$$

The estimate of the phase ϕ_0 for this adaptive measurement step corresponds to the phase with the maximum posterior probability $\max_\phi\{P(\phi|n, \beta)\}$. After this measurement step, the posterior distribution $P(\phi|n, \beta)$ becomes the prior distribution $P(\phi)$ for the subsequent adaptive measurement step, and this procedure is repeated for all L adaptive periods.

Optimization of the strategy for phase estimation requires the optimization of the displacement operation $\hat{D}(\beta_{\mathrm{opt}})$ in phase $\arg(\beta)$ and amplitude $|\beta|$ in each adaptive measurement step prior to photon detection. This optimization is achieved by maximizing a given objective function that depends on the prior $P(\phi)$ and the likelihood $\mathcal{L}(n|\phi, \beta)$ distributions, averaged over all possible detection results in the measurement step [see Sec. I of the Supplemental Material (SM) [43]]. While several objective functions can be used for this optimization [44], here we focus on two objective functions: the expected sharpness of the posterior distribution $\langle S(\beta, m)\rangle$ [14,21,45] and the mutual information $I(\beta, m)$ [38,46,47].

The expected sharpness of the posterior for a given measurement step with PNR(m) is

$$\langle S(\beta, m)\rangle = \sum_{n=0}^m P(n) \left| \int_0^{2\pi} e^{i\phi} P(\phi|n, \beta) d\phi \right|, \quad (2)$$

where $P(n) = \int \mathcal{L}(n|\phi, \beta)P(\phi)d\phi$ is the probability of detecting n photons. Therefore, the optimal displacement field β_{opt} maximizes $\langle S(\beta, m)\rangle$ over all possible detection results n in that step, resulting in maximal expected sharpness.

The expected mutual information for an adaptive measurement step can be written as [38,46,47]

$$I(\beta, m) = \sum_{n=0}^{m} \int_0^{2\pi} P(\phi, n|\beta) \log_2 \left[\frac{P(\phi, n|\beta)}{P(n)P(\phi)} \right] d\phi. \quad (3)$$

Here $P(\phi, n|\beta) = P(\phi|n, \beta)P(n) = P(n|\phi, \beta)P(\phi)$ is the joint probability distribution for n and ϕ given β, and the optimal displacement β_{opt} maximizes the mutual information over all possible detection results n.

Recursive application of this procedure over the L adaptive steps over the single-shot measurement yields a complete history of photon detections $\{n\}_L$ and optimal displacements $\{\beta\}_L$, which are used to calculate the final phase estimate:

$$\hat{\phi} = \arg[\langle e^{i\phi} \rangle] = \arg \left[\int_0^{2\pi} e^{i\phi} P(\phi|\{n\}_L, \{\beta\}_L) d\phi \right]. \quad (4)$$

Here $P(\phi|\{n\}_L, \{\beta\}_L)$ is the final reconstructed posterior distribution given the complete measurement history. After many (N) repetitions of the measurement with initially random relative phases between the input state and the initial displacement field, the variance of the distribution of phase estimates from Eq. (4) can be described by the Holevo variance for cyclic variables [29,48,49], which is bounded by the CRLB:

$$\text{Var}[\hat{\phi}] = \frac{1}{|\langle e^{i\hat{\phi}} \rangle|^2} - 1 \geq \frac{1}{4|\alpha|^2}, \quad (5)$$

where $|\langle e^{i\hat{\phi}} \rangle| = |\sum_{j=1}^{N} e^{i\hat{\phi}_j}|/N$ is the sharpness of the distribution of final estimates $\{\hat{\phi}\}_N$.

Figure 1(b) shows the Holevo variance multiplied by the quantum Fisher information (QFI $= 4|\alpha|^2$) for the optimized non-Gaussian estimation strategies with $L = 30$, PNR(3), and ideal detection efficiency. These results are obtained through Monte Carlo simulations using different objective functions: sharpness (blue) and mutual information (orange). Bold lines represent the average of five Monte Carlo simulations each with $N = 10^3$ randomly sampled initial phases ϕ_0, and the shaded regions represent one standard deviation. Also shown is the CRLB (solid black) for coherent states given by $1/4|\alpha|^2$ and the lower bound on the variance of an ideal heterodyne measurement (dashed black) given by $1/2|\alpha|^2$.

The gray line shows the expected variance of a strategy without an optimized displacement field, which highlights the advantages of optimized strategies. A nonoptimized strategy uses a fixed local oscillator (LO) amplitude $|\beta| = |\alpha|$, but with a LO phase that is adaptively set to the current phase estimate $\hat{\phi}$. In such a strategy, the LO is always attempting to displace the estimated state $|\alpha e^{i\hat{\phi}}\rangle$ to the vacuum state. Also shown for reference is the expected performance of the adaptive homodyne strategy, termed

"Mark II," proposed in [29,30](green line). We observe that both optimized non-Gaussian estimation strategies enable estimation with variances below the heterodyne limit from $|\alpha|^2 \approx 2.5$ to $> 10^3$ and reach a minimum of 1.13 times the CRLB at $|\alpha|^2 \approx 50$.

We note that the increase in variance of the estimation strategies at large $|\alpha|^2$ can be mitigated by increasing the number of adaptive steps L or the number resolution of the detector, as shown in Fig. 1(b), which we have further investigated [44]. We observe that at $|\alpha|^2 = 10^3$ the strategies can reach variances within 10% from the CRLB. In particular, the strategy optimizing mutual information with $L = 200$ and PNR(3) achieves the smallest variance of 1.003 ± 0.03 times the CRLB. [50] These studies suggest that optimized photon counting measurements can achieve variances within 1% from the CRLB at finite mean photon numbers. Figure 1(b) also shows that for both Gaussian and non-Gaussian measurement strategies, as $|\alpha|^2$ decreases there is an increase in the phase variance. This behavior is consistent with that of the canonical phase measurement [30,44,49].

We find that during the initial steps of the measurement (small k), the optimal LO values for each objective function are in general very different due to the non-Gaussian photon counting statistics. However, as the measurement progresses, for large k the posterior phase distributions for both strategies approach a Gaussian distribution with small variance, which has two notable consequences. First, in this limit the optimal LO parameters for both strategies asymptote to the same values, and these values make the classical Fisher information for displaced photon counting equal to the QFI (see Sec. II of the SM [43] for details). This result is consistent with the theoretical work in [47] showing that Bayesian experimental designs that adaptively optimize the mutual information (or the variance of the posterior distribution and by extension the sharpness) are asymptotically efficient, thus reaching the CRLB [51]. Second, in this limit of phase distributions with small variances, maximization of the mutual information is analogous to maximization of the Fisher information due to the connection between the mutual information and Fisher information via the relative entropy, i.e., Kullback-Leibler divergence [52,53]. In contrast, maximization of the expected sharpness will yield minimal variance [48]. This observation can be interpreted as each strategy attempting to reach the CRLB but through different approaches: either maximizing the Fisher information or minimizing the phase variance. These asymptotic findings are consistent with our results for both strategies, which show similar performances in the overall variance.

Experimental demonstration.—Figure 2 shows the diagram for the experimental demonstration of the optimized estimation strategies based on an interferometric setup. A helium-neon laser and an acousto-optic modulator pulsed at 1 kHz prepare coherent state pulses that enter the interferometer in a 50/50 beam splitter. An attenuator and a
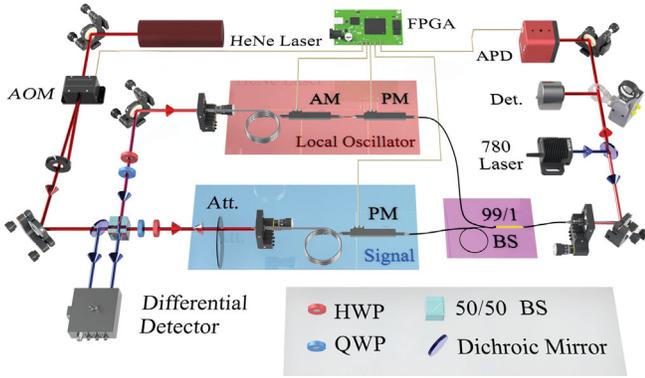
FIG. 2. Experimental setup for the adaptive displaced photon counting measurement (see text for details). HeNe: helium-neon laser, AOM: acousto-optic modulator, Att.: attenuator, PM: phase modulator, AM: amplitude modulator, BS: beam splitter, APD: avalanche photodiode; FPGA: field-programmable gate array.

phase modulator prepare the input coherent state with a fixed mean photon number $|\alpha|^2$ and a phase ϕ_0 to be estimated. The optimized displacement β_{opt} is prepared in an LO field with an amplitude modulator and a second phase modulator. We use a 99/1 beam splitter to implement the displacement operation $\hat{D}(\beta_{opt})|\alpha\rangle$. The photons in the displaced state are detected by an avalanche photodiode that acts as a PNR detector with finite photon number resolution [41,42]. A field-programmable gate array (FPGA) collects the detection events and implements the phase estimation strategies described above. This FPGA realizes in real time the Bayesian estimation procedure during each adaptive measurement, as described in Sec. I of the SM [43], and optimizes the displacement $\hat{D}(\beta_{opt})$ conditioned on photon detections. The optimization protocol in the FPGA uses a Gaussian approximation for steps $k \geq 9$ in which the phase distributions $P(\phi)$ are approximated as Gaussian distributions (see Sec. II in the SM [43]). This approximation allows us to demonstrate strategies with large numbers of adaptive steps, $L = 30$. At the end of an estimation measurement, the FPGA transfers the history of detections $\{n\}_L$, the phase estimates $\{\hat{\phi}\}_L$, and the optimized displacements $\{\beta_{opt}\}_L$ to a computer for processing and final estimation based on Eq. (4). We note that while the optimization protocol uses a Gaussian approximation, the recursive Bayesian procedure uses the complete phase distribution $P(\phi)$.

For experimental convenience, instead of randomly choosing the initial phase ϕ_0 and a fixed initial displacement phase $\theta_\beta^{(0)} = \arg(\beta) = 0$, we fix $\phi_0 = \pi$ and start at a random phase $\theta_\beta^{(0)}$. These two situations are equivalent, since the initial relative phase $\theta_\beta^{(0)} - \phi_0$ is random in both cases (see Sec. III of the SM [43]). Our experimental implementation has a $\approx 66\%$ duty cycle, uses a laser at 780 nm to stabilize the interferometric setup [41,42,54,55], and achieves an overall system efficiency $\eta = 0.70$ (avalanche photodiode efficiency $\eta_{APD} \approx 0.84$), interference

visibility $\xi = 0.997$, and dark count rate $\nu = 140/s$. We demonstrate estimation strategies for mean photon numbers from $|\alpha|^2 = 1$ to $|\alpha|^2 = 10^3$, with PNR(3) and $L = 30$ adaptive measurement steps that are each 20 μs long.

Results.—Figure 3 shows, as a function of $|\alpha|^2$, the Holevo variance of the experimental results for single-shot phase estimation of coherent states multiplied by the QFI $(4|\alpha|^2)$. Data points are the results for strategies maximizing the average sharpness $\langle S(\beta, m) \rangle$ (blue) and the mutual information $I(\beta, m)$ (MI, orange). The points and error bars represent the average variance and one standard deviation over five different runs of the experiment, respectively, each of which is the variance of $N = 10^4$ independent experiments with random initial relative phases $\theta_\beta^{(0)} - \phi_0$. The CRLB (solid red line), the heterodyne limit (dashed red line), and these bounds adjusted to the efficiency $\eta = 0.70$ (solid and dashed black lines) are included for reference. The purple (green) solid line is the expected Holevo variance for a strategy maximizing the sharpness (mutual information) obtained from the average of five Monte Carlo simulations of the experiment, each with 10^3 samples. These simulations take into account the experimental imperfections and the effects of limited resolution, precision, and bandwidth of the FPGA on the implementation of
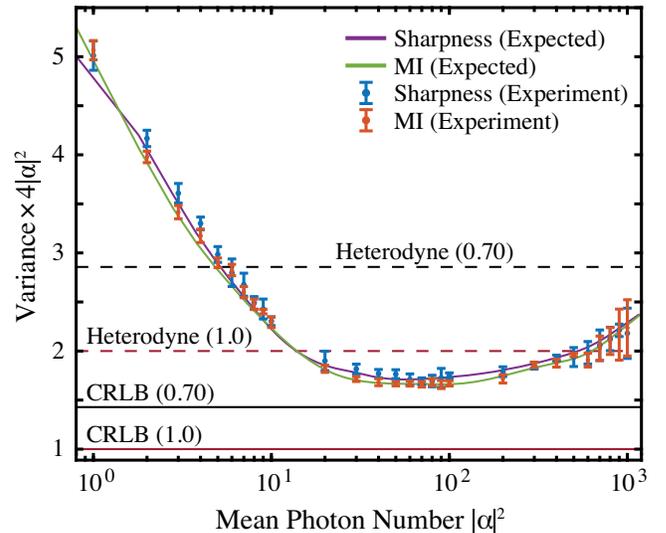

FIG. 3. Experimental results. Experimentally obtained Holevo variance multiplied by the QFI for coherent states $(4|\alpha|^2)$ for estimation strategies maximizing the sharpness and the mutual information as blue and orange points, respectively. Also shown are the expected variances (solid lines) accounting for experimental parameters. Included are the CRLB for coherent states $1/4|\alpha|^2$ (solid red line) and the ideal heterodyne limit $1/2|\alpha|^2$ (dashed red line), together with these bounds adjusted to the efficiency of our implementation $\eta = 0.70$ (solid and dashed black lines). Note that both strategies outperform the ideal heterodyne bound from $|\alpha|^2 \approx 20$ to ≈ 600 and the adjusted bound from $|\alpha|^2 \approx 6$ to $> 10^3$. Furthermore, both estimation strategies achieve a minimum of ≈ 1.68 times the CRLB (1.18 times adjusted) at $|\alpha|^2 \approx 70$.

the strategies. The increase of the error bars for large $|\alpha|^2$ in the experimental results is due to the increased sensitivity of the strategies to environmental noise at larger mean photon numbers. As the variance is small in this regime, technical noise causing small fluctuations or instabilities in the relative phase of the input state and LO may become non-negligible compared to the measured variance.

We observe that the optimized non-Gaussian estimation strategies surpass the ideal heterodyne bound from $|\alpha|^2 \approx 20$ to 600, and the adjusted ($\eta = 0.70$) heterodyne bound from $|\alpha|^2 \approx 6$ to $> 10^3$. Furthermore, in both strategies the variance reaches a minimum at $|\alpha|^2 \approx 70$ of less than 1.7 times the CRLB and less than 1.2 times adjusted for $\eta = 0.70$. These results show that adaptive non-Gaussian measurements enable highly robust phase estimation below the heterodyne limit without correction for detection inefficiencies. We note that increasing the number of adaptive steps and photon number resolution [44] could allow for increasing the measurement sensitivity at higher $|\alpha|^2$.

Conclusion.—We propose and demonstrate optimized non-Gaussian estimation strategies for single-shot measurements of an unknown phase of optical coherent states with sensitivities surpassing the heterodyne limit and approaching the CRLB. These strategies use optimized displacement operations, single-photon counting with finite photon number resolution, and a moderate number of adaptive steps with fast feedback. Our demonstration of the single-shot estimation strategies uses fast processing for optimization of the displacement operations in real time conditioned on the detection history as the measurement progresses throughout the single optical mode. We show that our experimental demonstration surpasses the ideal heterodyne limit for a wide range of optical powers without correcting for detection efficiency and achieves variances of less than 1.2 times the equivalent CRLB adjusted to our system efficiency. This is, to our knowledge, the most sensitive single-shot measurement of an unknown phase encoded in optical coherent states to date.

We expect that these optimized measurements for phase estimation of coherent states can be used to enhance the performance of schemes based on measurement backaction for the cooling of mechanical oscillators [32,33], the preparation of spin squeezed states [34], and applications in waveform and force detection [36]. Moreover, optimized non-Gaussian measurements can potentially be used for surpassing the limits of Gaussian approaches for estimating a time-varying phase [56–59] and maximizing information transmission in optical communications [55,60], as well as for phase estimation when quantum resources are employed [31].

*Corresponding author.
fbecerra@unm.edu

[1] C. Xu, L. Zhang, S. Huang, T. Ma, F. Liu, H. Yonezawa, Y. Zhang, and M. Xiao, Photonics Res. 7, A14 (2019).

[2] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, Rev. Mod. Phys. 90, 035006 (2018).

[3] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics 5, 222 (2011).

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

[5] R. Demkowicz-Dobrzaski, M. Jarzyna, and J. Koodyski, Prog. Opt. 60, 345 (2015).

[6] M. G. Genoni, S. Olivares, and M. G. A. Paris, Phys. Rev. Lett. 106, 153603 (2011).

[7] C. Lee, C. Oh, H. Jeong, C. Rockstuhl, and S.-Y. Lee, arXiv:190706349.

[8] M. Bradshaw, P. K. Lam, and S. M. Assad, Phys. Rev. A 97, 012106 (2018).

[9] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, arXiv:2003.05821v1.

[10] M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. 106, 090401 (2011).

[11] K. Iwasawa, K. Makino, H. Yonezawa, M. Tsang, A. Davidovic, E. Huntington, and A. Furusawa, Phys. Rev. Lett. 111, 163602 (2013).

[12] M. Tsang, New J. Phys. 15, 073005 (2013).

[13] 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).

[14] Z. Huang, K. R. Motes, P. M. Anisimov, J. P. Dowling, and D. W. Berry, Phys. Rev. A 95, 053837 (2017).

[15] B. E. Anderson, P. Gupta, B. L. Schmittberger, T. Horrom, C. Hermann-Avigliano, K. M. Jones, and P. D. Lett, Optica 4, 752 (2017).

[16] B. E. Anderson, B. L. Schmittberger, P. Gupta, K. M. Jones, and P. D. Lett, Phys. Rev. A 95, 063843 (2017).

[17] S. Izumi, M. Takeoka, K. Wakui, M. Fujiwara, K. Ema, and M. Sasaki, Phys. Rev. A 94, 033842 (2016).

[18] S. Slussarenko, M. M. Weston, H. M. Chrzanowski, L. K. Shalm, V. B. Verma, S. W. Nam, and G. J. Pryde, Nat. Photonics 11, 700 (2017).

[19] S. Daryanoosh, S. Slussarenko, D. W. Berry, H. M. Wiseman, and G. J. Pryde, Nat. Commun. 9, 4606 (2018).

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

[21] A. Hentschel and B. C. Sanders, Phys. Rev. Lett. 104, 063603 (2010).

[22] Z. Hou, R.-J. Wang, J.-F. Tang, H. Yuan, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Phys. Rev. Lett. 123, 040501 (2019).

[23] W. Larson and B. E. A. Saleh, Phys. Rev. A 96, 042110 (2017).

[24] A. Lumino, E. Polino, A. S. Rab, G. Milani, N. Spagnolo, N. Wiebe, and F. Sciarrino, Phys. Rev. Applied 10, 044033 (2018).

[25] K. Zheng, H. Xu, A. Zhang, X. Ning, and L. Zhang, Quantum Inf. Process. 18, 329 (2019).

[26] D. W. Berry, H. M. Wiseman, and J. K. Breslin, Phys. Rev. A 63, 053804 (2001).

[27] R. Demkowicz-Dobrzaski, J. Koodyski, and M. Gu, Nat. Commun. **3**, 1063 (2012).

[28] H. M. Wiseman, Phys. Rev. Lett. **75**, 4587 (1995).

[29] H. M. Wiseman and R. B. Killip, Phys. Rev. A **56**, 944 (1997).

[30] H. M. Wiseman and R. B. Killip, Phys. Rev. A **57**, 2169 (1998).

[31] A. A. Berni, T. Gehring, B. M. Nielsen, V. Hndchen, M. G. A. Paris, and U. L. Andersen, Nat. Photonics **9**, 577 (2015).

[32] M. R. Vanner, J. Hofer, G. D. Cole, and M. Aspelmeyer, Nat. Commun. **4**, 2295 (2013).

[33] S. Seidelin, Y. Le Coq, and K. Mølmer, Phys. Rev. A **100**, 013828 (2019).

[34] I. Bouchoule and K. Mølmer, Phys. Rev. A **66**, 043811 (2002).

[35] J. Z. J. Aasi and J. Abadie, Nat. Photonics **7**, 613 (2013).

[36] M. Tsang and R. Nair, Phys. Rev. A **86**, 042115 (2012).

[37] G. M. D'Ariano, M. G. A. Paris, and R. Seno, Phys. Rev. A **54**, 4495 (1996).

[38] I. Bargatin, Phys. Rev. A **72**, 022316 (2005).

[39] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, Phys. Rev. Lett. **89**, 133602 (2002).

[40] L. S. Martin, W. P. Livingston, S. Hacohen-Gourgy, H. M. Wiseman, and I. Siddiqi, arXiv:1906.07274.

[41] F. E. Becerra, J. Fan, and A. Migdall, Nat. Photonics **9** (2015).

[42] M. T. DiMario and F. E. Becerra, Phys. Rev. Lett. **121**, 023603 (2018).

[43] See Supplemental Material http://link.aps.org/supplemental/10.1103/PhysRevLett.125.120505 as for a discussion of the algorithm used in the FPGA to implement the adaptive measurement strategy; the Gaussian approximation of the prior distribution used in the algorithm for the experiment; and the preparation of the initial relative phase between the input state and the local oscillator.

[44] M. Rodriguez-Garcia, M. T. DiMario, P. Barberis-Blostein, and F. E. Becerra, Adaptive single-shot phase estimation strategies for coherent states (to be published).

[45] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).

[46] W. Rzadkowski and R. Demkowicz-Dobrzanski, Phys. Rev. A **96**, 032319 (2017).

[47] L. Paninski, Neural Comput. **17**, 1480 (2005).

[48] 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).

[49] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Edizioni della Normale, Pisa, 2011).

[50] We note that some datasets for this point lay below the CRLB, but this effect is due to finite sampling ($N = 10^3$) in the simulations and a larger N would reduce this effect.

[51] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications and Signal Processing (Wiley-Interscience, New York, 2006).

[52] S. Kullback, *Information Theory and Statistics*, 1st ed. (Dover Publications, New York, 1968), p. 11.

[53] N. Brunel and J. Nadal, Neural Comput. **10**, 1731 (1998).

[54] F. E. Becerra, J. Fan, G. Baumgartner, J. Goldhar, J. T. Kosloski, and A. Migdall, Nat. Photonics **7**, 147 (2013).

[55] A. R. Ferdinand, M. T. DiMario, and F. E. Becerra, npj Quantum Inf. **3**, 43 (2017).

[56] L. Zhang, K. Zheng, F. Liu, W. Zhao, L. Tang, H. Yonezawa, L. Zhang, Y. Zhang, and M. Xiao, Opt. Express **27**, 2327 (2019).

[57] T. A. Wheatley, D. W. Berry, H. Yonezawa, D. Nakane, H. Arao, D. T. Pope, T. C. Ralph, H. M. Wiseman, A. Furusawa, and E. H. Huntington, Phys. Rev. Lett. **104**, 093601 (2010).

[58] D. T. Pope, H. M. Wiseman, and N. K. Langford, Phys. Rev. A **70**, 043812 (2004).

[59] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph *et al.*, Science **337**, 1514 (2012).

[60] J. Lee, S.-W. Ji, J. Park, and H. Nha, Phys. Rev. A **93**, 050302(R) (2016).