PHYSICAL REVIEW A **70**, 043812 (2004)

Adaptive phase estimation is more accurate than nonadaptive phase estimation for continuous beams of light

D. T. Pope,[1,*] H. M. Wiseman,[1,2,†] and N. K. Langford[3,‡]
[1]*Centre for Quantum Dynamics, School of Science, Griffith University, Brisbane 4111, Queensland, Australia*
[2]*Centre for Quantum Computer Technology, Griffith University, Brisbane 4111, Australia*
[3]*Centre for Quantum Computer Technology, School of Physical Sciences, University of Queensland, Brisbane 4072, Queensland, Australia*

(Received 14 May 2004; published 19 October 2004)

We consider the task of estimating the randomly fluctuating phase of a continuous-wave beam of light. Using the theory of quantum parameter estimation, we show that this can be done more accurately when feedback is used (adaptive phase estimation) than by *any* scheme not involving feedback (nonadaptive phase estimation) in which the beam is measured as it arrives at the detector. Such schemes not involving feedback include all those based on heterodyne detection or instantaneous canonical phase measurements. We also demonstrate that the superior accuracy of adaptive phase estimation is present in a regime conducive to observing it experimentally.

DOI: 10.1103/PhysRevA.70.043812 PACS number(s): 42.50.Dv, 42.50.Lc, 03.67.Hk

I. INTRODUCTION

Phase is a physical property found in both classical and quantum electromagnetic (EM) fields. For classical EM fields comprising a single mode, it can be determined exactly via measuring two orthogonal quadratures or components of such fields. This, however, is not the case for single-mode EM fields in quantum mechanics. Estimates of the phases of such fields are necessarily imperfect due to intrinsic quantum noise in measurements of noncommuting observables such as quadratures. Given this limitation, quantum *phase estimation*, the process of estimating the phase of a quantum-mechanical EM field as accurately as possible, is nontrivial.

In addition to being nontrivial, phase estimation in quantum mechanics is interesting for a number of reasons. First, at some time in the future it may be practical to encode and send information in the phase of a single electromagnetic field mode at or near the ultimate quantum limit—the upper limit permitted by quantum mechanics [1–3]. In such a scenario, the more accurately a receiver could estimate phase the more information could be sent. Second, it also may be useful in interferometric gravity-wave detection. Third, phase estimation is interesting as it is an instance of quantum parameter estimation [4,5], an increasingly experimentally accessible field concerned with estimating parameters of quantum states as well as possible in the face of unavoidable quantum noise.

Phase can be estimated via two broad approaches, *nonadaptive* phase estimation and *adaptive* phase estimation [5]. In nonadaptive phase estimation, which is the conventional approach, we measure an EM field via a single fixed measurement that remains constant over time. In adaptive phase estimation, however, the measurement is continually adjusted in an attempt to maximize its accuracy at each moment in time. This is done by changing or adapting it based on earlier measurement results. For both EM-field pulses and also continuous EM beams, it has been shown that adaptive phase estimation is more accurate than (at least) many instances of the conventional nonadaptive approach [5–8].

In this paper we consider the problem of estimating the randomly fluctuating phase of a continuous-wave (cw) EM field (EM beam) as introduced in Ref. [8]. We show that this can be done more accurately using adaptive phase estimation than via any nonadaptive phase estimation scheme in which the field is measured in real time (that is, as it arrives at the detector). We also show that this improved accuracy exists for fields with small to moderate photon fluxes. These are our two main results. The latter is significant, first, as a *theoretical* difference between the accuracies of adaptive and nonadaptive phase estimation is most readily seen *experimentally* in fields with small to moderate photon fluxes. Second, in a communication scenario in which a receiver is trying to extract information encoded in the phase of an EM field by a distant sender, it is likely that the receiver will be making measurements on fields with small to moderate photon fluxes due to attentuation [9]. In the course of arriving at the two results, we present a theoretical technique for estimating phase that may be applicable to a range of problems. Our results build upon earlier work [5–8] and further demonstrate the superiority of adaptive schemes over conventional nonadaptive ones for the important task of phase estimation.

This paper is structured as follows. In Sec. II, we review the mathematical tools used throughout. They are Bayes' rule, the Kushner-Stratonovitch equation, and the Zakai equation. Next, Sec. III presents the phase estimation schemes considered, some of which are adaptive and some of which are nonadaptive. In Sec. IV, we compare the accuracies of the schemes in the steady-state regime, showing that each of the adaptive schemes is more accurate than all of the nonadaptive schemes. Finally, in Sec. V we discuss our results.

*Electronic address: d.pope@griffith.edu.au
†Electronic address: h.wiseman@griffith.edu.au
‡Electronic address: langford@physics.uq.edu.au

Before proceeding further, we first review existing work on adaptive phase estimation. As previously stated, the conventional method for estimating the phase of an EM field is via nonadaptive phase estimation. For a single-mode EM-field *pulse* in the coherent state $|\beta\rangle$, where $\beta \in \mathbb{C}$, the most widely known method [2,9] of estimating the phase ϕ $[=\arg(\beta)]$ uses a nonadaptive detection technique called heterodyne detection [2,10–14]. This involves mixing the pulse, which we call the *signal pulse*, with an intense local oscillator of phase $\Phi = \Phi_0 + \Delta t$ at a 50:50 beam splitter. Here Δ is a detuning, t denotes time, and Φ_0 is the phase at $t=0$. The difference between the photocurrents in the beam splitter's two output ports is proportional to the quadrature phase amplitude $X_\Phi = ae^{-i\Phi} + a^\dagger e^{i\Phi}$, where a and $a^\dagger$ are creation and annihilation operators for the signal pulse. Assuming that $\Delta \gg \Gamma$, where Γ is the signal pulse's spectral width, all quadratures are rapidly measured and thus, for all practical purposes, heterodyne detection instantaneously measures the complex photocurrent I_c containing equal information about the observables $X_{\Phi=0}$ and $X_{\Phi=\pi/2}$. Once the signal pulse has been measured, ϕ can then be estimated from an appropriate functional of all the recorded currents. For large values of $|\beta|$ this approach leads to an estimate with a variance of $1/(2|\beta|^2)$ [6]. Half of this is nonfundamental and results from excess noise introduced by heterodyne detection due to the fact that it measures two noncommuting quadratures. This excess contribution to the variance can also be thought of as arising from the fact that heterodyne detection measures all quadratures equally. Because of this, it sometimes measures the so-called amplitude quadrature ($X_{\Phi=\phi}$) which contains no information about ϕ.

A second type of EM field for which phase estimation has been considered is a *continuous* EM beam. In particular, Ref. [8] considered such estimation for a continuous beam in a coherent state with phase ϕ that randomly fluctuated in time as a Wiener process [15]. This paper found that one particular nonadaptive phase estimation scheme estimated ϕ with a variance of $1/\sqrt{2N}$ in the steady-state regime for $N \gg 1$. Here, N is the beam's photon flux in an amount of time equal to its coherence time (which is set by the time scale of the fluctuations in ϕ).

Though nonadaptive phase estimation using heterodyne detection yields a reasonable estimate of ϕ for both a single EM-field pulse and a continuous EM beam, this quantity can be more accurately estimated via adaptive techiques [5–8]. For a single pulse of light, again in the coherent state $|\beta\rangle$, this can be done by measuring the field using adaptive homodyne detection. Nonadaptive homodyne detection is identical to heterodyne detection except that the local oscillator has the same frequency as the pulse's mean frequency so that Φ is constant [16]. It is made adaptive by varying Φ so as to try to measure the so-called phase quadrature. This is the quadrature for which $\Phi = \phi + \pi/2$, and, moreover, the one that minimizes the measurement's excess uncertainty, below that of heterodyne detection. To try to measure the phase quadrature we use the results of previous measurements to obtain $\hat{\phi}_{\rm fb}(t)$, an estimate for $\phi(t)$. This is then fed back to the local oscillator and Φ is set to $\Phi(t) = \hat{\phi}_{\rm fb}(t) + \pi/2$ in an attempt to "home in" on the phase quadrature. Figure 1 shows a schematic diagram of the apparatus implementing this scheme. When $|\beta|$ is large, it leads to a variance in our estimate of $1/(4|\beta|^2)$ [6], which is only half as large as that of the nonadaptive scheme discussed above. Furthermore, this improved accuracy has been seen experimentally [9].

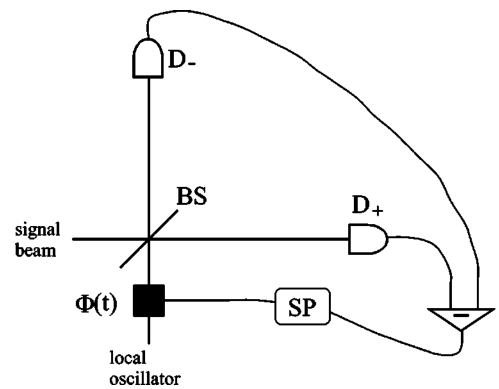

FIG. 1. Schematic diagram of the measurement setup for adaptive homodyne-based phase-estimation schemes. The symbol BS denotes the 50:50 beam splitter; D_- and D_+ are photon counters for which the difference in the number of photons they detect is found and then fed back to the local oscillator's phase. A signal processor is denoted by SP.

For the continuous EM beam with a randomly fluctuating phase considered earlier, it is known that a particular adaptive scheme is more accurate than one particular nonadaptive one [8]. But is it also more accurate than the best possible nonadaptive scheme? One of the main results of this paper is to show, in Sec. IV, that in the steady-state regime adaptive phase estimation is more accurate than any nonadaptive estimation scheme in which the EM field is measured in real time, even one involving a canonical phase measurement [17]. In addition, we show that the improved accuracy of adaptive phase estimation persists for $N \ll 1$.

II. BACKGROUND THEORY

A. What is phase?

Within quantum mechanics, the term "phase" has multiple meanings [18,17]. In this paper, however, it refers to a single concept which we now state. The electric field of a *classical* single-mode EM-field pulse incident on an ideal photodetector is, in the vicinity of this detector,

$$\vec{E}(t) = \sqrt{\frac{2\hbar\omega}{\epsilon_0 \mathcal{A} c}}\vec{e}(|\alpha|e^{-i(\omega t - \phi_{\rm cl})} + \text{c.c.}). \tag{2.1}$$

Here, t denotes time, ω is the field's angular frequency, ϵ_0 denotes the permittivity of free space, $\mathcal{A}$ is the transverse area over which the field is spread, $\vec{e}$ represents a unit vector denoting the field's direction, $|\alpha|$ is a complex amplitude with dimensions of time$^{-1/2}$, c denotes the speed of light, and c.c. represents a complex conjugate. Given this, we define $\phi_{\rm cl}$ to be this field's phase. Similarly, the phase of a quantum-mechanical single-mode EM-field pulse is defined to be the quantum-mechanical analog of $\phi_{\rm cl}$, which we denote by ϕ. For instance, the phase of the coherent state $||\beta|e^{i\phi}\rangle$ is de-

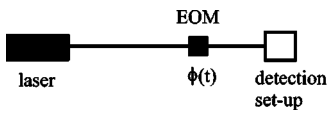

FIG. 2. Schematic diagram of the physical scenario considered. First, an idealized cw laser outputs a continuous beam of light which is then incident on an electro-optical modulator (EOM). The EOM imprints phases on segments of the beam which are then incident on the detection setup on the right.

fined to be ϕ which is a parameter and not an observable. In particular, it is not the observable associated with the Pegg-Barnett phase operator [19] which is also called phase but which does not have a well-defined value for the state $||\beta|e^{i\phi}\rangle$.

B. Continuous EM beam

The scenario that we consider throughout this paper centers around a continuous EM beam [8] known as the *signal beam*. This beam is the output of an idealized laser, and so can be described by a coherent state with complex amplitude α. The mean photon flux $|\alpha|^2$ is constant. However, the beam's phase $\phi(t)$ fluctuates randomly such that, again in the vicinity of the detector,

$$\frac{d\phi}{dt} = \sqrt{\kappa}\xi(t). \tag{2.2}$$

Here κ is a noise strength and ξ is real Gaussian white noise defined by

$$\langle \xi(t)\xi(t')\rangle = \delta(t-t'). \tag{2.3}$$

In practice, this fluctuation could be achieved via an electro-optical modulator [20] that "imprints" a fluctuating phase on each segment of the beam. These phase fluctuations give the beam a linewidth of κ, so that $N=|\alpha|^2/\kappa$ is the number of photons in the coherence time (see Fig. 2).

In the continuous EM beam scenario, we measure the signal beam via either homodyne or heterodyne detection. For homodyne detection, the photocurrent I_r measured in the interval dt is given by

$$I_r dt = 2\eta\,\mathrm{Re}(|\alpha|e^{i(\phi-\Phi)})dt + \sqrt{\eta}\,dW. \tag{2.4}$$

Here dW is a real Wiener increment, η is the detector's efficiency (which is its probability of detecting an incident photon), and Φ is the local oscillator phase. In contrast, heterodyne detection simultaneously measures the quadratures $X_{\Phi=0}$ and $X_{\Phi=\pi/2}$. An alternate way of doing this is to first split the signal beam at a 50:50 beam splitter and then to measure $X_{\Phi=0}$ at one output and $X_{\Phi=\pi/2}$ at the other. Assuming perfect detectors, each photodetector measures, on average, half of the beam's photons and thus the quantum efficiency of each measurement is $\eta=1/2$. Representing both outcomes in terms of a single complex quantity, we obtain

$$I_c dt = |\alpha|e^{i\phi}dt + dW_c, \tag{2.5}$$

where dWc is a complex Wiener increment defined by the correlations $\langle dW_c dW_c^*\rangle = dt$ and $\langle dW_c dW_c\rangle = 0$.

C. Nonadaptive and adaptive phase estimation

In a number of the phase-estimation schemes we consider, ϕ is estimated using the theory of quantum parameter estimation [4,5]. This process involves two steps. First, Bayes' rule is used to obtain a differential equation with respect to time for $P(\phi)$, the probability distribution encoding our knowledge of ϕ, which we then solve. Bayes' rule updates our knowledge of some unknown parameter given the measurement result M. For the situations we consider, it is

$$P(\phi|M) = \frac{P(\phi)P(M|\phi)}{P(M)}, \tag{2.6}$$

where $P(x|y)$ denotes the probability of x given y. The second step in the process of estimating ϕ via quantum parameter estimation is to use $P(\phi)$ to calculate our estimate of ϕ, which we denote by $\hat{\phi}(t)$.

To explain in more detail the first step of generating and solving a differential equation for $P(\phi)$, we begin by observing that in Eq. (2.6) the term $P(M)$ is a normalization factor that ensures the normalization of $P(\phi|M)$. This can be seen by realizing that we can write $P(M)$ as

$$P(M) = \int_{\phi=\phi_0}^{\phi_0+2\pi} P(\phi)P(M|\phi)d\phi, \tag{2.7}$$

where ϕ_0 is an arbitrary lower limit. It follows from this that upon replacing $P(M)$ in Eq. (2.6) by another function of M that is independent of ϕ we obtain a quasi-Bayes' rule that updates an *unnormalized* "probability" distribution for ϕ that we label $\tilde{P}(\phi)$ [4]. We choose to replace $P(M)$ by $P(M)_{|\alpha|=0}$, where $P(M)_{|\alpha|=0}$ is the probability of measuring the result M given $|\alpha|=0$, and so Eq. (2.6) becomes

$$\tilde{P}(\phi|M) = \frac{\tilde{P}(\phi)P(M|\phi)}{P(M)_{|\alpha|=0}}. \tag{2.8}$$

The function $P(M)_{|\alpha|=0}$ was chosen as it corresponds to considering the measurement result M in the denominator to be Gaussian white noise which, in turn, simplifies Eq. (2.6). Furthermore, it yields a linear evolution equation for $\tilde{P}(\phi)$. This is in contrast to the nonlinear one for $P(\phi)$ that would have been obtained had $P(M)$ not been replaced.

The next step in obtaining and solving a differential equation for $P(\phi)$ is to transform Eq. (2.8) into the form

$$d\tilde{P}(\phi) = [f(\phi)g(M) + \text{c.c.}]\tilde{P}(\phi)dt, \tag{2.9}$$

where $f(\phi)$ and $g(M)$ are functions whose nature depends upon $P(M|\phi)$ and $P(M)_{|\alpha|=0}$, by neglecting terms of order dt^2 or higher. This equation is known as a *Zakai equation* [21]. To obtain the desired differential equation for $P(\phi)$ with respect to time from Eq. (2.9) we normalize $\tilde{P}(\phi)$ using a known procedure detailed in Appendix A. This leads to the following differential equation for $P(\phi)$:

$$dP(\phi) = |\alpha|[(e^{i\phi} - \langle e^{i\phi}\rangle_{P(\phi)})P(\phi)\zeta(t) + \text{c.c.}]dt, \tag{2.10}$$

where ζ is either real or complex Gaussian white noise depending on the nature of M. This is known as a *Kushner-Stratonovitch (KS) equation* [22].

Thus far, we have only considered the evolution of $P(\phi)$ due to our measurement of the signal beam. However, there is also its evolution resulting from the diffusion described by Eq. (2.2). In the absence of measurement, this diffusive evolution leads to $P(\phi)$ being a Gaussian distribution centered on $\phi(t=0)$ with variance κt. A straightforward calculation shows that the evolution equation for this distribution in this case is the Fokker-Planck equation

$$dP(\phi) = \frac{\kappa}{2}\frac{\partial^2 P(\phi)}{\partial \phi^2}dt. \tag{2.11}$$

Adding the effects of phase diffusion to Eq. (2.20) leads to the final KS equation

$$dP = \frac{\kappa}{2}\frac{\partial^2 P}{d\phi^2}dt + |\alpha|[(e^{i\phi} - \langle e^{i\phi}\rangle_{P(\phi)}P\zeta(t) + \text{c.c.}]dt. \tag{2.12}$$

Solving this equation we obtain $P(\phi)$.

As stated at the start of this subsection, the second step in estimating $\phi(t)$ via quantum parameter estimation is to calculate the optimal estimate for $\phi(t)$ from $P(\phi)$. This is defined to be the estimate with the following two properties.

(1) It has the smallest possible average error as measured by the Holevo variance [23].

(2) It is such that $\langle \exp[i(\phi-\hat{\phi})]\rangle_{I,\xi} \in \mathbb{R}_+$. Here $\langle\cdots\rangle_{I,\xi}$ is an average over I and ξ, where I is either I_c or I_r depending on the measurement scheme.

The Holevo variance is a measure of statistical spread suitable for any cyclical variable x and is given by

$$V^H(x) = |\langle e^{ix}\rangle|^{-2} - 1. \tag{2.13}$$

For such variables, it is superior to the standard variance σ^2 as the latter can be ill defined. To illustrate this problem, observe that ϕ has the range $[\phi_0, \phi_0+2\pi)$, where ϕ_0 is usually chosen to be either $-\pi$ or 0. As a result, depending on our choice of $\phi_0, \sigma^2(\phi)$ can take different values for a single distribution. The reason for the second property is to rule out estimates with small Holevo variances but which are systematically biased and hence do not estimate ϕ accurately.

The optimal estimate we wish to calculate is given by

$$\hat{\phi}(t) = \arg[\langle \exp(i\phi)\rangle_{P(\phi)}], \tag{2.14}$$

where $\langle\cdots\rangle_{P(\phi)}$ denotes an average over $P(\phi)$. While the estimate $\langle\phi\rangle_{P(\phi)}$ is a more obvious choice for the optimal estimate of $\phi(t)$, it sometimes estimates $\phi(t)$ poorly due to the fact that $\phi(t)$ is cyclical. This occurs, for instance, when $P(\phi)$ is centered near ϕ_0. It is important to realize that the estimate in Eq. (2.14) is not optimal in an absolute sense. Rather, it is the best estimate of ϕ given that we have chosen to minimize the "cost function" $V^H(\phi)$.

It is interesting to note that the approach to estimating $\phi(t)$ outlined above differs from that in other work on phase estimation [5–8]. These other papers generated estimates based on intuitive, partially justified mathematical functions and, as a consequence, their estimates were sometimes suboptimal. In contrast, a number of this paper's phase-estimation schemes use quantum parameter estimation which leads to optimal estimates for $\phi(t)$ [at least according to the cost or error function $V^H(\phi)$].

To illustrate our method of obtaining $\hat{\phi}(t)$ via quantum parameter estimation, we now demonstrate its application in the case of measuring the signal beam via heterodyne detection. (Its use in the other cases we consider is very similar.) For this type of detection, Bayes' rule is

$$P(\phi|I_c) = \frac{P(\phi)P(I_c|\phi)}{P(I_c)}. \tag{2.15}$$

Replacing the normalization constant $P(I_c)$ by $P(I_c)_{|\alpha|=0}$ leads to the quasi-Bayes' rule

$$\tilde{P}(\phi|I_c) = \frac{\tilde{P}(\phi)P(I_c|\phi)}{P(I_c)_{|\alpha|=0}}. \tag{2.16}$$

Equation (2.5) tells us that the real and imaginary parts of I_c are Gaussian random variables with variances of $1/(2dt)$ and, respectively, means of $|\alpha|\cos\phi$ and $|\alpha|\sin\phi$. From this it follows that

$$\tilde{P}(I_c|\phi) = \sqrt{\frac{dt}{\pi}}\exp(-dt\{[\text{Re}(I_c) - |\alpha|\cos\phi]^2 + [\text{Im}(I_c) - |\alpha|\sin\phi]^2\}) \tag{2.17}$$

while

$$P(I_c)_{|\alpha|=0} = \sqrt{\frac{dt}{\pi}}\exp\left(-\frac{dt}{2}I_c^* I_c\right). \tag{2.18}$$

Substituting the expressions on the right-hand side of Eqs. (2.17) and (2.18) into Eq. (2.16) and neglecting terms of order dt^2 or higher leads to the Zakai equation

$$d\tilde{P}(\phi) = |\alpha|(e^{i\phi}I_c + \text{c.c.})\tilde{P}(\phi)dt. \tag{2.19}$$

Normalizing $\tilde{P}$ via the known procedure detailed in Appendix A, from Eq. (2.19) we obtain the Kushner-Stratonovitch equation [22]

$$dP(\phi) = |\alpha|[(e^{i\phi} - \langle e^{i\phi}\rangle_{P(\phi)})P(\phi)\zeta(t) + \text{c.c.}]dt, \tag{2.20}$$

where ζ is complex Gaussian white noise ($\zeta=I_c - |\alpha|\langle e^{i\phi}\rangle_{P(\phi)}$ and is the so-called observation or measurement noise [24]). Incorporating the effects of phase diffusion, we arrive at

TABLE I. Summary of phase estimates.

Name of measurement scheme	$\hat{\phi}$	$d\Phi/dt$	Type of detection
Canonical	$\arg[\langle\exp(i\phi)\rangle_{P(\phi)}]$	N/A	Canonical
Optimal heterodyne based	$\arg[\langle\exp(i\phi)\rangle_{P(\phi)}]$	Δ	Heterodyne
BW heterodyne based	$\arg(A_t)$	Δ	Heterodyne
BW adaptive	$\arg(A_t+\chi B_t A_t^*)$	$\sqrt{\kappa}I_r$	Homodyne
Semioptimal adaptive	$\arg[\langle\exp(i\phi)\rangle_{P(\phi)}]$	$d\hat{\phi}/dt$	Homodyne
Simple adaptive	$\arg(A_t)$	$\sqrt{\kappa}I_r$	Homodyne

$$dP(\phi) = \frac{\kappa}{2}\frac{\partial^2 P(\phi)}{d\phi^2}dt + |\alpha|[(e^{i\phi} - \langle e^{i\phi}\rangle_{P(\phi)})P(\phi)\zeta(t) + \text{c.c.}]dt. \tag{2.21}$$

Note that this equation has been previously derived, albeit for a different (but related) physical system via a different method [25]. It is also interesting to realize that we could have obtained Eq. (2.21) via beginning with Eq. (2.6), substituting into it expressions for $P(I_c|\phi)$ and $P(I_c)$, and performing some algebra while neglecting terms of order dt^2 or higher. Although this method is conceptually simpler than the one we used, it involves a more challenging calculation. To complete the process of determining $\hat{\phi}$, once we have obtained Eq. (2.21), we solve it and then use $P(\phi)$ to calculate $\hat{\phi}(t)$ via Eq. (2.14).

III. PHASE-ESTIMATION SCHEMES

In this paper we compare the accuracies of a number of nonadaptive and adaptive phase estimation schemes for an EM beam. Prior to doing so, however, we outline the schemes considered, detailing nonadaptive and adaptive schemes in turn. These are summarized in Table I.

A. Nonadaptive schemes

1. Berry-Wiseman heterodyne-based scheme

In the *Berry-Wiseman (BW) heterodyne-based* phase-estimation scheme [8] the signal beam is measured via heterodyne detection. The phase estimate at time t, $\hat{\phi}(t)$, is then calculated from the measurement record up to t. Specifically, it is

$$\hat{\phi}(t) = \arg(A_t), \tag{3.1}$$

where A_t can be written as

$$A_t = \int_{u=-\infty}^{t} du\, e^{\chi(u-t)} I_c(u), \tag{3.2}$$

where χ is a scaling parameter. More specifically, χ scales the weight $\exp[-\chi(u-t)]$ given to each current I_u. While this estimate may not seem intuitive, it was chosen as an analogous estimate for the single-shot scenario was known to be accurate [6]. Moreover, Ref. [8] showed that, for large N, $\arg A_t$ was an accurate estimate for a continuous EM beam when χ was set to $\chi = 2|\alpha|\sqrt{\kappa}$.

2. Optimal heterodyne-based scheme

In this scheme, the signal beam is measured via heterodyne detection and then, following the calculation in Sec. II, quantum parameter estimation is used to obtain the KS equation Eq. (2.20). This is then solved and its solution used to obtain $\hat{\phi}(t)$ in accordance with Eq. (2.14).

3. Canonical scheme

The *canonical* phase estimation scheme involves making a canonical phase measurement [17] on the signal beam at each instant in time and then taking $\hat{\phi}(t)$ to be its outcome. Naively, it might be thought that this scheme would be more accurate than any other as a canonical measurement, or so it is thought, is the most accurate measurement of phase one can make. Results in Sec. IV show, however, that this is not the case (for reasons explained in Sec. V).

B. Adaptive schemes

1. Simple adaptive scheme

In the *simple adaptive* phase-estimation scheme [8] we measure the signal beam via adaptive homodyne detection and then estimate $\phi(t)$ to be

$$\hat{\phi}(t) = \arg(A_t), \tag{3.3}$$

where here

$$A_t = \int_{u=-\infty}^{t} du\, e^{\chi(u-t)} e^{i\Phi} I_r(u). \tag{3.4}$$

We also adapt the homodyne measurement, setting the local oscillator's phase to $\Phi(t) = \hat{\phi}(t) + \pi/2$. From this it follows [8] that it is updated such that its rate of change with time is

$$\frac{\partial \Phi}{\partial t} = \sqrt{\kappa} I_r(t). \tag{3.5}$$

This equation follows from letting $\chi = 2|\alpha|\sqrt{\kappa}$ in Eq. (3.4) which is known to be optimal for large N [8]. One of the reasons the simple adaptive scheme was considered in Ref. [8] was that the fact that for large N it was known to be optimal. In Section IV we show that it also performs well for small to moderate values of N.

2. Berry-Wiseman adaptive scheme

The *Berry-Wiseman adaptive* phase-estimation scheme involves measuring the signal beam via adaptive homodyne detection. The phase estimate at time $t, \hat{\phi}(t)$, is then a function of two functionals of all measurement results up to time t. Specifically, it is

$$\hat{\phi}(t) = \arg(A_t + \chi B_t A_t^*), \tag{3.6}$$

where A_t is as defined in Eq. (3.2) and B_t is

$$B_t = \int_{u=-\infty}^{t} du\, e^{\chi(u-t)} e^{2i\Phi(u)}. \tag{3.7}$$

As for the BW heterodyne-based scheme, this estimate was chosen as an analogous estimate for the single-shot case was known to be accurate [6]. Furthermore, Ref. [8] showed that it was accurate for large N, for $\chi = 2|\alpha|\sqrt{\kappa}$.

3. Semioptimal adaptive scheme

In the *semioptimal adaptive* scheme for phase estimation, we assume it is optimal to always measure the signal beam's phase quadrature and thus, as in the other adaptive schemes, set $\Phi(t) = \hat{\phi}(t) + \pi/2$. The reason we use the label "semioptimal adaptive" is that, while we use quantum parameter estimation in determining $\hat{\phi}$, we are not certain that it is always best to attempt to measure the phase quadrature. Perhaps one could obtain a more accurate estimate by occasionally trying to measure the amplitude quadrature, for example.

IV. RESULTS

To compare the accuracies of the estimates introduced in Sec. III, we now calculate their average errors as measured by the Holevo variance V^H of the difference between the actual phase ϕ and our estimate $\hat{\phi}$. Typically, this quantity fluctuates for some time before settling down to a fixed steady-state value. Intuitively, this occurs as a balance arises (on average) between the information we gain about ϕ from a new photocurrent measurement and that we lose due to ϕ's phase diffusion over the measurement's duration. We choose this steady-state value of $V^H(\phi - \hat{\phi})$, denoted by $V^H_{\rm SS}$, as our measure of the efficacy of our phase-estimation schemes and hence numerically determine it for all of them for a range of N values. We also obtain analytic expressions for it for some schemes for both large and small values of N.

From the definition of the Holevo variance in Eq. (2.13), $V^H(\phi - \hat{\phi})$ is given by

$$V^H(\phi - \hat{\phi}) = |\langle e^{i(\phi-\hat{\phi})}\rangle_{\xi,I}|^{-2} - 1, \tag{4.1}$$

where the average $\langle \cdots \rangle_{\xi,I}$ is a stochastic average over ξ and I. To calculate this quantity for our three estimates generated via parameter estimation, we first use the fact that

$$\langle e^{i(\phi-\hat{\phi})}\rangle_{\xi,I} = \langle |\langle e^{i\phi}\rangle_{P(\phi)}|\rangle_I \tag{4.2}$$

to express $V^H(\phi - \hat{\phi})$ as

$$V^H(\phi - \hat{\phi}) = \langle |\langle e^{i\phi}\rangle_{P(\phi)}|\rangle_I^{-2} - 1. \tag{4.3}$$

A demonstration of Eq. (4.2) is given in Appendix B. After arriving at Eq. (4.3), we then use the ergodic theorem within this equation to replace the ensemble average $\langle |\langle e^{i\phi}\rangle_{P(\phi)}|\rangle_I$ *in the steady state* by the temporal average

$$\frac{1}{t_f - t_0^{\rm SS}} \int_{t=t_0^{\rm SS}}^{t_f} dt |\langle e^{i\hat{\phi}}(t)\rangle_{P(\phi)}|, \tag{4.4}$$

where $t_0^{\rm SS}$ is the time at which the steady-state regime begins and t_f is the final time we consider ($t_f \gg t_0^{\rm SS}$). This allows us to determine $V^H_{\rm SS}$ through simulating just a single stochastic trajectory.

Upon calculating $V^H_{\rm SS}$, a number of trends are apparent. The first of these concerns the proximity of $\hat{\phi}$ to ϕ in the simple adaptive scheme. For large N, the initial estimate $\hat{\phi}(t=0)$ for this scheme is usually some distance from the actual phase $\phi(t=0)$. Then, as we gain more and more information via measurement and postprocessing, $\hat{\phi}$ homes in on ϕ during an initial period of transience. After this it locks onto ϕ, staying close to ϕ as it continues to fluctuate a little. This pattern of behavior is illustrated in Fig. 3(a). It is anticipated that all the schemes considered behave similarly, although we did not explicitly verify this. For small values of $N, \hat{\phi}$ never locks onto ϕ but instead continues to fluctuate in its vicinity with a magnitude that increases with decreasing N, as highlighted in Fig. 3(b).

A second trend in our results concerns the size of the interval within which we are fairly certain that ϕ lies at any moment in time. This is measured by the Holevo variance $|\langle e^{i\phi}\rangle_{P(\phi)}|^{-2} - 1$ which can be thought of as a measure of our lack of confidence in $\hat{\phi}$. For large N, this quantity, at least for the schemes based on parameter estimation, only fluctuates over time by a small amount once the initial transience ends. This behavior is illustrated in Fig. 4(a). It can be explained by realizing that when N is large we are in a linear regime in the sense that the measured photocurrent I_r or I_c is a *linear* function of the actual phase ϕ. For instance, for homodyne detection we have $I_r dt = 2\eta|\alpha|(\phi - \hat{\phi}) + \sqrt{\eta}\, dW$. It is a characteristic trait of such linear systems that our level of confidence (and hence also our lack of confidence) in any estimate of a system parameter is constant in the steady state [22]. For small N, however, $|\langle e^{i\phi}\rangle_{P(\phi)}|^{-2} - 1$ fluctuates appreciably for all t (for the schemes based on parameter estimation), as shown in Fig. 4(b).

A. Nonadaptive schemes

1. Berry-Wiseman heterodyne-based scheme

Previous work [8] has calculated $V^H_{\rm SS}$ for the BW heterodyne-based scheme for a range of N values. These results are plotted in Fig. 5. For large N, it is known [8] that the scheme has a steady-state error of $V^H_{\rm SS} \simeq 1/\sqrt{2N}$.

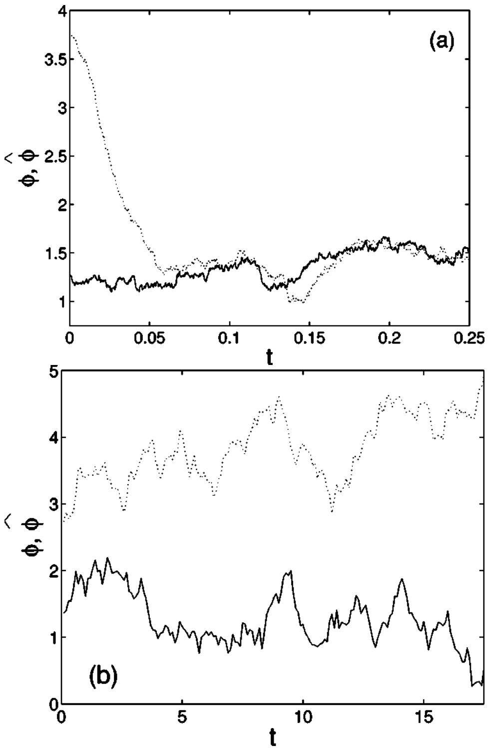

FIG. 3. Graphs showing typical variations of the actual phase ϕ (solid line) and our estimate $\hat{\phi}$ (dotted line) versus time t scaled by κ for the simple-adaptive phase-estimation scheme for (a) a large photon flux ($N=1000$) and (b) a small one ($N=0.1$). In (a), $\hat{\phi}$ initially homes in on ϕ, before locking onto it. In (b), the low photon flux means we gain so little information from our measurements that $\hat{\phi}$ never locks onto ϕ. Both ϕ and $\hat{\phi}$ are dimensionless, as is t.

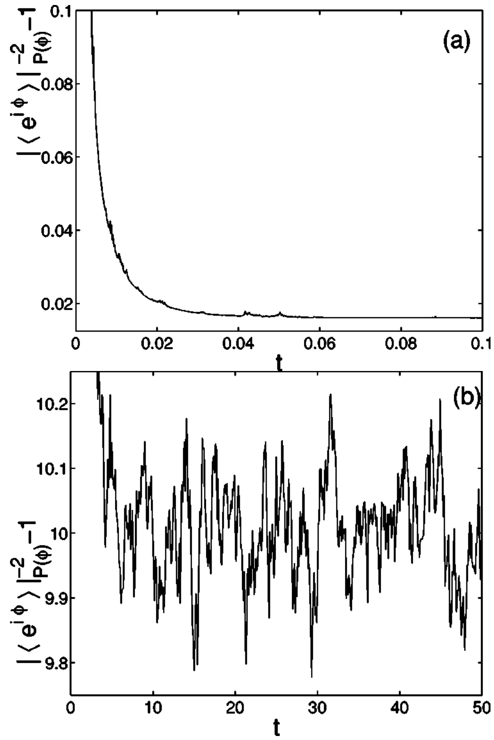

FIG. 4. Graphs showing typical variations of our lack of confidence in $\hat{\phi}$ as measured by $|\langle e^{i\phi}\rangle_{P(\phi)}|^{-2}-1$ (dimensionless) versus time t (dimensionless) scaled by κ for the semioptimal adaptive phase-estimation scheme for (a) a large photon flux ($N=1000$) and (b) a small one ($N=0.1$).

2. Optimal heterodyne-based scheme

For the optimal heterodyne-based phase-estimation scheme, $V_{\rm SS}^H$ was calculated by determining the temporal average in expression (4.4) and then using Eq. (4.3) to find $V_{\rm SS}^H$. This was done by, first, expressing $P(\phi)$ in Eq. (2.21) as the following discrete Fourier series:

$$P(\phi)=\sum_{j=-\infty}^{\infty} b_j \exp(ij\phi), \tag{4.5}$$

where $b_j \in \mathbb{C}$ and $b_{-j}=b_j^*$. Next, the resulting equation was transformed into Fourier space to produce the following coupled differential equations:

$$\dot{b}_j=-\frac{\kappa j^2 b_j}{2}+|\alpha|\zeta b_{j-1}+|\alpha|\zeta^* b_{j+1}-4\pi b_j|\alpha|\mathrm{Re}(\zeta^* b_1). \tag{4.6}$$

These were then numerically solved by considering only b_j's for which $|j|$ was less than some finite bound that increased with N. Next, $\langle e^{i\phi}\rangle_{P(\phi)}(t)$ was determined by exploiting the fact that it is a function of just one Fourier coefficient ($|b_1|$). Finally, we averaged over numerous steady-state values of $\langle e^{i\phi}\rangle_{P(\phi)}(t)$ to obtain expression (4.4) and thus $V_{\rm SS}^H$. The results generated are plotted in Fig. 5. Analytic results were also found for large and small N which are $V_{\rm SS}^H \simeq 4/(\pi N)$ (small N) and $V_{\rm SS}^H \simeq 1/\sqrt{2N}$ (large N).

Our analytic result for $V_{\rm SS}^H$ for small N was obtained by first realizing that when $N \ll 1$ heterodyne measurements on the signal beam yield little information about ϕ and thus $P(\phi)$ is broad. This means that, in contrast, $P(\phi)$'s Fourier transform is narrow and, more specifically, that the following relations hold (on average): $|b_0| \gg |b_1| \gg |b_2| \ldots$. Because of this, we can neglect Fourier coefficients for which $|j|>1$ in

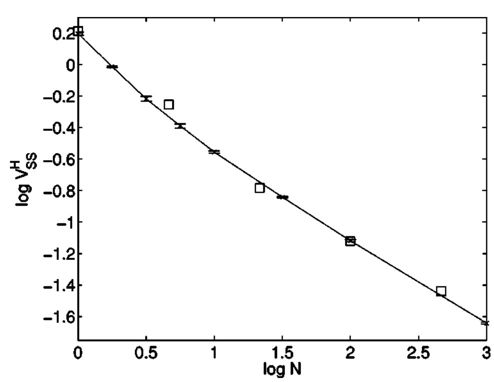

FIG. 5. Log-log plot (to base 10) of the steady-state Holevo variance $V_{\rm SS}^H$ versus photon flux N for the BW heterodyne-based nonadaptive (squares) and optimal heterodyne nonadaptive (solid line) phase estimation schemes. Both $V_{\rm SS}^H$ and N are dimensionless.

Eq. (4.6). Upon doing this, and also neglecting terms containing $|b_1|^2$ (as $|b_1|^2 \ll |b_1|$), we are left with just the following equation for b_1:

$$\frac{\partial b_1}{\partial t} = -\frac{\kappa b_1}{2} - \frac{|\alpha|\zeta}{2\pi}. \quad (4.7)$$

Note that $b_0(t)=1/(2\pi)$ [as can be determined from the normalization condition $\int_{\phi=\phi_0}^{2\pi+\phi_0} P(\phi)d\phi=1$]. Solving Eq. (4.7) we find that in the steady state b_1 is a complex Gaussian random variable with mean zero and a variance of $N/(8\pi^2)$ in both its real and imaginary parts.

To calculate $V_{\rm SS}^H$ from b_1 we first note that

$$|\langle e^{i\phi}\rangle_{P(\phi)}| = \left| \int_{\phi=\phi_0}^{\phi_0+2\pi} d\phi\, e^{i\phi}P(\phi)\right|. \quad (4.8)$$

Substituting $P(\phi)=\Sigma_{j=-\infty}^{\infty} b_j \exp(ij\phi)$ into the right-hand side of this equation yields

$$|\langle e^{i\phi}\rangle_{P(\phi)}| = 2\pi|b_1|. \quad (4.9)$$

From this it follows that the equation

$$V_{\rm SS}^H = \langle |\langle e^{i\phi}\rangle_{P(\phi)}|\rangle_I^{-2} - 1 \quad (4.10)$$

simplifies to

$$V_{\rm SS}^H = (2\pi\langle |b_1|\rangle_\zeta)^{-2} - 1. \quad (4.11)$$

Given that $\langle |b_1|\rangle_\zeta \simeq \sqrt{N}/(4\sqrt{\pi})$ we obtain $V_{\rm SS}^H \simeq 4/(\pi N) - 1$. Neglecting the second term (as this produces a more accurate approximation) yields

$$V_{\rm SS}^H \simeq 4/(\pi N). \quad (4.12)$$

The large-N approximation for $V_{\rm SS}^H$ for the optimal heterodyne scheme was obtained by replacing the exponents in Eq. (2.21) by a linear approximation and then assuming that $P(\phi)$ was Gaussian. Differential equations with respect to time for the mean and variance of this Gaussian were then constructed and solved to obtain the standard variance of $\hat{\phi}$ in the steady state which, for large N, is approximately equal to $V_{\rm SS}^H$.

The expression $(e^{i\phi} - \langle e^{i\phi}\rangle_{P(\phi)})\zeta$ in Eq. (2.21) can be reexpressed as

$$(e^{i(\phi-\hat{\phi})} - \langle e^{i(\phi-\hat{\phi})}\rangle_{P(\phi)})e^{i\hat{\phi}}\zeta. \quad (4.13)$$

When $N \gg 1$, the large photon fluxes present in the signal beam mean that our measurements yield a great deal of information about ϕ and hence that $\hat{\phi}$ is a highly accurate estimate. As a result, $e^{i(\phi-\hat{\phi})} \simeq 1$ and thus we can linearize expression (4.13) as follows:

$$(e^{i(\phi-\hat{\phi})} - \langle e^{i(\phi-\hat{\phi})}\rangle_{P(\phi)})e^{i\hat{\phi}}\zeta \simeq i(\phi - \langle\phi\rangle_{P(\phi)})e^{i\hat{\phi}}\zeta. \quad (4.14)$$

The expression $e^{i\hat{\phi}}\zeta$ behaves as complex Gaussian white noise and hence we denote it as ζ'. Substituting the above results into Eq. (2.21), we obtain

$$dP(\phi) = \frac{\kappa}{2}\frac{\partial^2 P}{d\phi^2}dt - 2|\alpha|[i(\phi - \langle\phi\rangle_{P(\phi)})\mathrm{Re}(\zeta')]dt. \quad (4.15)$$

To solve this equation, we assume that $P(\phi)$ is Gaussian and thus that it can be expressed as

$$P(\phi) = \frac{\exp[-(\phi-\mu_P)^2/(2\sigma_P^2)]}{\sqrt{2\pi}\sigma_P}, \quad (4.16)$$

where μ_P and σ_P^2 are, respectively, P's mean and variance. Generating differential equations for μ_P and σ_P, we obtain

$$d\sigma_P^2 = d\langle\phi^2\rangle_{P(\phi)} - d(\langle\phi\rangle_{P(\phi)}^2) = d\langle\phi^2\rangle_{P(\phi)} - 2\langle\phi\rangle_{P(\phi)}d\langle\phi\rangle_{P(\phi)} - (d\langle\phi\rangle_{P(\phi)})^2 \quad (4.17)$$

and

$$d\mu_P = -2|\alpha|\sigma_P^2 \mathrm{Re}(\zeta')dt. \quad (4.18)$$

Solving these yields

$$\sigma_P^2(t) = \frac{1}{\sqrt{2N}}\frac{\exp(2\sqrt{2}|\alpha|^2 t/\sqrt{N}) + 1}{\exp(2\sqrt{2}|\alpha|^2 t/\sqrt{N}) - 1}. \quad (4.19)$$

In the limit of $t \to \infty$ this reduces to

$$\sigma_{P\ \rm SS}^2 \simeq V_{\rm SS}^H \simeq \frac{1}{\sqrt{2N}}. \quad (4.20)$$

Interestingly, this result is the same as that obtained in [8]. This shows that the BW heterodyne-based scheme, which was designed for large N, is indeed optimal in this regime.

3. Canonical scheme

For the canonical phase-estimation scheme, $\hat{\phi}(t)$ was calculated via quantum parameter estimation using the method in Sec. II C. For this scheme, Bayes' rule is

$$P(\phi|\theta) = \frac{P(\phi)P(\theta|\phi)}{P(\theta)}, \quad (4.21)$$

where θ is the measured phase. As a canonical phase measurement is a projective measurement of the Pegg-Barnett phase observable [19], the probability of it yielding the result θ is $(2\pi)^{-1}$ times the square of the norm of the measured state's projection onto the (unnormalized) phase eigenstate $|\theta\rangle = \Sigma_{n=0}^{\infty} e^{in\theta}|n\rangle$. Thus, for the coherent states we consider, to first order in $\sqrt{dt}$,

$$P(\theta|\phi) = \frac{1}{2\pi}|\langle\alpha\sqrt{dt}|\theta\rangle|^2 = \frac{1}{2\pi}[1 + 2|\alpha|\sqrt{dt}\cos(\theta-\phi)] \quad (4.22)$$

and thus

$$P(\theta)_{|\alpha|=0} = (2\pi)^{-1}. \quad (4.23)$$

Substituting the expressions on the right-hand sides of Eqs. (4.22) and (4.23) into Eq. (4.21) leads to the following Zakai equation:

$$d\tilde{P}(\phi) = \frac{|\alpha|}{\sqrt{dt}}(e^{i(\phi-\theta)} + \text{c.c.})\tilde{P}(\phi)dt. \tag{4.24}$$

Using the known correspondence detailed in Appendix A, this, in turn, leads to the KS equation

$$dP(\phi) = 2 \times \text{Re}\left[|\alpha|\left((e^{i(\phi)} - \langle e^{i\phi}\rangle_{P(\phi)})\frac{e^{-i\theta}}{\sqrt{dt}}P(\phi)\right)dt\right]. \tag{4.25}$$

Letting $e^{-i\theta}/\sqrt{dt}=f$, we find that $\langle f\rangle = \langle f^2\rangle = 0$ (at least when we average over any finite time interval) and $\langle ff^*\rangle = 1/dt$ from which it follows that f is complex Gaussian white noise. Given this, Eq. (4.25) reduces to Eq. (2.20), the KS equation obtained for the optimal heterodyne-based phase-estimation scheme. As a result, the canonical scheme shares the same accuracy as this other scheme and so shares the same results for V_{SS}^H. This surprising result is explained in Sec. V.

4. Comparison

As can be seen from Fig. 5, when $N \lesssim 10$, the optimal heterodyne-based phase-estimation scheme is slightly more accurate than the BW heterodyne-based one. For larger values of N, however, we see that both schemes seem to be equally accurate. (At approximately $N=10^{1.25}$, the BW heterodyne-based scheme appears to be more accurate, but this is due to numerical errors, primarily in the BW heterodyne-based result.) The first of these features illustrates that while the BW heterodyne-based scheme is close to optimal for $N \lesssim 10$, ϕ can be estimated more accurately using parameter estimation in this regime. The latter fact is particularly significant as this regime is the one in which an experimental realization could most readily be performed, as discussed in more detail in Sec. V. The second feature highlights that the BW heterodyne-based scheme is optimal for $N \gtrsim 10$ which is unsurprising as it was designed for large N [8].

B. Adaptive schemes

1. Simple adaptive scheme

For the simple adaptive phase-estimation scheme, the Holevo variance in the steady state was calculated by simulating the evolution of $\phi(t)$ via solving Eq. (2.2) and also simulating the measurement outcomes on the beam using Eq. (2.4) to obtain a numerical expression for $I_r(t)$ for a range of times. This allowed us to update $\hat{\phi}$ via

$$\frac{\partial \hat{\phi}}{\partial t} = \sqrt{\kappa}I_r(t) \tag{4.26}$$

and thus to determine $\phi(t)-\hat{\phi}(t)$, again for a range of times. The local-oscillator phase $\Phi(t)$ was then set to $\Phi(t)=\hat{\phi}(t)+\pi/2$. The steady-state Holevo variance V_{SS}^H was calculated from the difference $\phi(t)-\hat{\phi}(t)$.

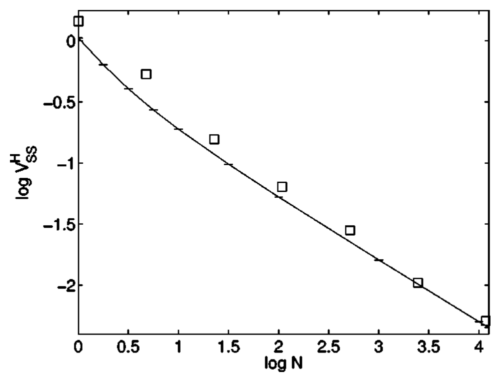

FIG. 6. Log-log plots (to base 10) of the logarithm of steady-state Holevo variance V_{SS}^H versus the photon flux N for the BW adaptive (squares) and the semioptimal adaptive (solid line) phase-estimation schemes. Both V_{SS}^H and N are dimensionless.

2. Berry-Wiseman adaptive scheme

For the BW adaptive scheme, Ref. [8] determined V_{SS}^H as a function of N and these results are shown in Fig. 6.

3. Semioptimal adaptive scheme

We derived $\hat{\phi}$ for the semioptimal adaptive scheme via quantum parameter estimation in the same manner as for the optimal heterodyne and canonical schemes. For this scheme, Bayes' rule is

$$P(\phi|I_r) = \frac{P(\phi)P(I_r|\phi)}{P(I_r)}. \tag{4.27}$$

Replacing the normalization constant $P(I_r)$ by $P(I_r)_{|\alpha|=0}$ yields the quasi-Bayes rule

$$\tilde{P}(\phi|I_r) = \frac{\tilde{P}(\phi)P(I_r|\phi)}{P(I_r)_{|\alpha|=0}}. \tag{4.28}$$

From Eq. (2.4) we know that I_r is a Gaussian random variable with variance $1/(dt)$ and mean $2|\alpha|\cos(\phi-\Phi)$ from which it follows that (for $\eta=1$)

$$P(\phi|I_r) = \sqrt{\frac{dt}{\pi}}\exp\{-dt[I_r - 2|\alpha|\cos(\phi-\Phi)]^2\} \tag{4.29}$$

and

$$P(I_r)_{|\alpha|=0} = \sqrt{\frac{dt}{\pi}}\exp(-dt\, I_r^2). \tag{4.30}$$

Substituting these two results into Eq. (4.28), we obtain the following Zakai equation:

$$d\tilde{P}(\phi) = |\alpha|(e^{i(\phi-\Phi)}I_r + \text{c.c.})\tilde{P}(\phi)dt. \tag{4.31}$$

Using the known correspondence detailed in Appendix A and including the effects of phase diffusion, Eq. (4.31) leads to the KS equation

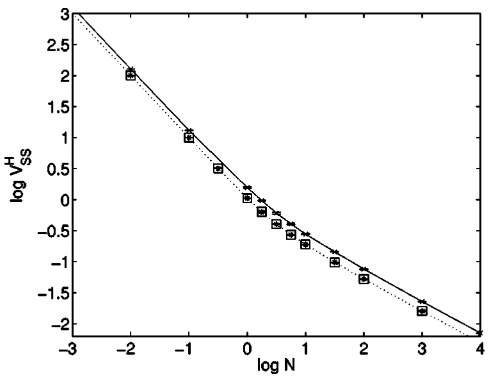

FIG. 7. Log-log plots (to base 10) of the logarithm of steady-state Holevo variance $V_{\rm SS}^H$ versus the photon flux N for the optimal heterodyne-based (solid line), semioptimal adaptive (squares), and simple adaptive phase-estimation schemes (asterisks). The large-N and small-N results lie upon the asymptotes derived for these regions. Both $V_{\rm SS}^H$ and N are dimensionless.

$$dP(\phi) = \frac{\kappa}{2}\frac{\partial^2 P(\phi)}{d\phi^2}dt + |\alpha|[(e^{i(\phi-\Phi)} - \langle e^{i(\phi-\Phi)}\rangle_{P(\phi)})P(\phi)\zeta'(t) + \text{c.c.}]dt, \qquad (4.32)$$

where ζ' is *real* Gaussian white noise given by $\zeta' = I_r - 2|\alpha|\langle\cos(\phi-\Phi)\rangle_{P(\phi)}$.

To obtain $V_{\rm SS}^H$ from Eq. (4.32) we applied the same method used for the optimal heterodyne-based scheme centered around decomposing $P(\phi)$ via the Fourier decomposition in Eq. (4.5). The results obtained are plotted in Fig. 6. In addition, for small and large N the following analytical results were found:

$$V_{\rm SS}^H \simeq \begin{cases} 1/(2\sqrt{N}) \quad (\text{large } N), & (4.33) \\ 1/N \quad (\text{small } N). & (4.34) \end{cases}$$

These results were obtained via calculations very similar to those used in Sec. IV A 2 to obtain the corresponding estimates for optimal heterodyne detection.

4. Comparison

Figures 6 and 7 display a number of interesting features which we now highlight. First, Fig. 7 shows that the semioptimal adaptive and simple adaptive schemes are equally accurate, as evidenced by the fact that they have identical $V_{\rm SS}^H$-versus-N plots. Second, Fig. 6 illustrates that the semioptimal adaptive scheme (and hence also the simple adaptive scheme) is more accurate than the BW adaptive one for all N values except when $N \gtrsim 10^{3.5}$. Third, Fig. 7 demonstrates that the semioptimal adaptive scheme is significantly more accurate than the optimal heterodyne-based or canonical schemes. Fourth, Fig. 7 also shows that adaptive phase estimation is more accurate than *any* nonadaptive phase-estimation scheme in which the field is measured in real time. The reason for this is the following. Assume that we measure the field nonadaptively in real time. By this we mean that we measure it via a continuous sequence of identical infinitesimal-time measurements and thus measure each spatial "segment" of the signal beam as it is incident on the detector. In this scenario, the best measurement we can make is a canonical phase measurement (as we must decide what to measure while knowing nothing about the phase). However, from Sec. III A 3, we know that estimating ϕ nonadaptively via such a measurement leads to an estimate only as accurate as that of the optimal heterodyne-based scheme. We also know that adaptive phase estimation is more accurate than this latter nonadaptive scheme in the cw scenario and hence it is also more accurate than the canonical nonadaptive scheme.

V. DISCUSSION

The results of Sec. IV display a number of interesting features which we now discuss. First, it might seem puzzling that the canonical phase-estimation scheme is only as accurate as the optimal heterodyne-based scheme and is not, instead, the most accurate scheme. Given that a canonical phase measurement is generally thought to be the best measurement of phase we can make, why isn't the canonical scheme the most accurate? The answer to this lies in the details of the scenario we consider. In the standard scenario in which we wish to estimate phase, we make a single phase measurement on a system for which we have no prior information about the phase. In this scenario, a canonical measurement is optimal. However, in the scenario we consider prior to making a measurement on the field at time $t \neq 0$, we already know something about ϕ, as evidenced by the fact that we possess a nontrivial probability distribution $P(\phi)$. This prior information can be exploited by measurements other than a canonical one to yield more information about phase than would a canonical measurement.

To understand the preceding point it may be helpful to consider the following example. Say we wish to determine as accurately as possible the phase of a system in a weak coherent state which we know to be either one of the two states $|\psi_\pm\rangle = |0\rangle + \gamma e^{\pm i\phi}|1\rangle$, where $\gamma \in \mathbb{R} \ll 1$, with equal probability. In this instance, because we already know something about ϕ, we can tailor the measurement in accordance with this prior knowledge and measure the $\Phi = \pi/2$ or Y quadrature to obtain slightly more information about ϕ than would a canonical measurement. Specifically, measuring the Y quadrature, we estimate ϕ correctly with probability $1/2 + 0.799\gamma\sin\phi$, while for a canonical measurement this probability is only $1/2 + 0.638\gamma\sin\phi$.

Another interesting feature related to Sec. IV's results concerns the main conclusion we drew from them, which was that adaptive phase estimation in the cw scenario is more accurate than any nonadaptive scheme in which the field is measured in real time. Although we were able to arrive at this result, we are uncertain if adaptive phase estimation is better than any nonadaptive scheme at all. This is because it is conceivable that there exists a nonadaptive scheme in which, instead of measuring the field in real time, we store up a portion of it over a period of time and then measure the accumulated field as a whole that is more accurate than adaptive phase estimation.

The results of Sec. IV also show that a simple adaptive scheme does as well as the semioptimal adaptive scheme.

Why does this relatively uncomplicated scheme do so well? One possibility is that the state of the beam we consider, being based on coherent states, is somewhat "simple." Perhaps, it does not allow us to fully utilize the power of the more complicated semioptimal adaptive scheme.

One final interesting feature of Sec. IV's results concerns the variation with N of the relative superiority of adaptive phase estimation over real-time nonadaptive phase estimation. This is measured by the ratio of the steady-state Holevo variances for the optimal heterodyne-based and the semioptimal adaptive schemes. For $N \ll 1$, this ratio is given by Eqs. (4.12) and (4.34) and is $4/\pi \simeq 1.27$ while for $N \gg 1$ it is $\sqrt{2} \simeq 1.41$. For intermediate N values, it lies in between these two extremes. Of particular importance is the fact that the gap is present for $N \simeq 1$. This is because this regime is the most fertile for experimental implementation as within it the errors we wish to see are not swamped by technical noise. It is also noteworthy that the small-N ratio of $4/\pi$ is significantly greater than the analogous ratio in Ref. [8], which was approximately 1.1, between the adaptive and nonadaptive estimates in this other paper.

Having discussed the results in Sec. IV, we now turn to three theoretical issues arising from our work. First, in this paper we have considered estimating the phase of an EM beam in a coherent state. However, other beams could be investigated as was done in Ref. [8] which looked at a so-called squeezed EM beam with a randomly fluctuating phase. That paper found that, for such a beam, adaptive phase estimation was more accurate than heterodyne-based nonadaptive phase estimation not just by a constant factor (as this paper has), but by a factor scaling with N. In particular, it found that for such a beam the steady-state Holevo variance of the error scaled as $N^{-2/3}$ in adaptive phase estimation but only as $N^{-1/2}$ in heterodyne-based nonadaptive phase estimation.

While this result for squeezed beams is interesting, the calculations behind it contained a number of deficiencies. First, Ref. [8] considered a beam with broadband squeezing, i.e., one that was squeezed at all frequencies, and thus the noise present in the beam had infinite energy. The parameter $N = |\alpha|^2/\kappa$ was finite, however, as it relates only to the energy carried by the mean field. Such a beam is unphysical and, furthermore, constitutes an inappropriate theoretical model for the problem considered, as we shall soon see. The second deficiency in the calculation was that it involved estimating $\phi(t)$ using only information about the beam's signal. This meant that information in the beam's noise was ignored. If such information had been used then, as the noise had infinite energy, we could have instantly determined ϕ by determining the relative sizes of the noise in different quadratures. Thus, the calculation in Ref. [8] ignored obtaining phase information from a potential source (the noise) and revolved around a model such that if we do consider this potential source, we find that we can instantly determine $\phi(t)$ with perfect accuracy, which is unrealistic. Because of these deficiencies, we feel that it is desirable to do additional calculations on squeezed beams. We anticipate that our "optimal" approach to obtaining phase estimates based on quantum parameter estimation may be useful in such calculations.

A second theoretical issue arising from this paper is the fact that throughout it we have assumed that the feedback present in the adaptive phase estimation schemes considered is instantaneous. That is, that it takes a zero amount of time to obtain an estimate of $\phi(t)$ and then transmit it to the local oscillator. This assumption, however, is unrealistic. In practice, this process would take a finite amount of time due to the fact that, for instance, a realistic signal processor would take a finite amount of time to calculate an estimate of $\phi(t)$ from information such as the measurement result at t. To give some examples, in the simple adaptive phase-estimation scheme a signal processor must calculate $\int \sqrt{\kappa} I_r(t) dt$ to obtain this estimate while in the semioptimal adaptive scheme it needs to update a probability distribution for ϕ in accordance with the KS equation Eq. (4.32) and then calculate arg $(\langle e^{i\phi(t)} \rangle_{P(\phi)})$. Previous work [6,26] has shown that the effect of such delays in feeding back estimates of $\phi(t)$ to the local oscillator is to increase the Holevo variance $V^H_{\rm SS}(\phi - \hat{\phi})$ of adaptive phase-estimation schemes. In turn, this means that they decrease the amount by which the simple adaptive and semioptimal adaptive schemes can estimate ϕ more accurately than can nonadaptive schemes. As we wish to maximize this amount, it seems that the simple adaptive scheme is preferable to the semioptimal adaptive one. While both schemes are equally accurate, the former calculates $\hat{\phi}$ via a simpler calculation which could be performed in less time. Consistent with this, it would be challenging to solve the KS equation Eq. (4.32) in a short enough time as to make an interesting experimental implementation of the theoretical work in this paper feasible.

The recent experimental implementation of adaptive phase estimation [9] used an almost identical estimate to that of the simple adaptive phase estimation scheme and involved a delay of approximately 0.1 μs. Interestingly, the main reason for this delay was *not* due to the signal processor having to perform a calculation. Instead, it was the speed at which a certain radio-frequency synthesizer in the experiment operated. Following on from this, as long as $t_{\rm delay}/t_{\rm coh} \ll 1$, where $t_{\rm delay}$ is the delay time in the feedback loop for some adaptive phase estimation scheme and $t_{\rm coh} (= \kappa^{-1})$ is ϕ's coherence time, $\phi(t)$ would not change appreciably in $t_{\rm delay}$ and thus a time delay in the feedback loop would not significantly increase the value of $V^H_{\rm SS}(\phi - \hat{\phi})$ for either the simple or the semioptimal adaptive scheme [27]. Assuming the time delay in Ref. [9], the above inequality could be satisfied by constraining κ such that $\kappa \ll 10^7 \text{ s}^{-1}$. This is achievable in practice as the electro-optical modulator in Fig. 2 can be changed slowly enough so as to satisfy the constraint $\kappa \ll 10^7 \text{ s}^{-1}$ without suffering appreciable decoherence. As a result, the presence of a realistic time delay does not seem to make it impossible to see the theoretical superiority of adaptive phase estimation.

One final theoretical issue arising from our work is the following. Throughout the paper, it was assumed that $|\alpha|$ was known precisely. However, even if we know only that $|\alpha| \geq a$, where $a \in \mathbb{R}$, we can still do at least as well as when we know that it equals a. This follows on from work by Stockton *et al.* [28] (Sec. V). Knowing $|\alpha|$ precisely, we have, for the simple adaptive (and semioptimal adaptive) schemes,

$$\chi_{\rm opt} = 2\sqrt{\kappa}|\alpha|. \tag{5.1}$$

If we know only that $|\alpha| \geq a$ we can set χ equal to

$$\chi = 2\sqrt{\kappa a}. \tag{5.2}$$

For $N \gg 1$, this leads to [8]

$$V_{\rm SS}^{H} = 2\sqrt{\kappa a}/(8\alpha^2) + \sqrt{\kappa}/(2a) \tag{5.3}$$

$$\simeq \sqrt{\kappa}/(2a). \tag{5.4}$$

That is, we can estimate ϕ at least as well as we can assuming we know that $|\alpha|$ is exactly the minimum known value.

VI. CONCLUSION

Quantum phase estimation and, in particular, Bayes' rule were used to find optimally accurate phase estimates and to show that, for a continuous EM beam with a randomly fluctuating phase, adaptive phase estimation is more accurate than any nonadaptive phase-estimation scheme in which the field is measured in real time. Although it is more accurate for all photon fluxes it is, in particular, more accurate for such beams possessing small to moderate photon fluxes. This is important as this is the regime in which experiments would have the greatest chance of confirming any theoretical difference between the two types of phase-estimation schemes.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council. D.T.P. would like to thank Dr. Dominic Berry, Dr. Kurt Jacobs, and Neil Oxtoby for their assistance.

APPENDIX A

This section details the known correspondence between a Zakai equation of the form

$$d\tilde{P} = (XI + \text{c.c.})\tilde{P}\, dt \tag{A1}$$

and the KS equation

$$dP = [(X - \langle X\rangle_P)(I - \langle I\rangle_P) + \text{c.c.}]P\, dt. \tag{A2}$$

To obtain Eq. (A2) from Eq. (A1), we begin with the identity

$$P(\phi) + dP(\phi) = \frac{\tilde{P} + d\tilde{P}}{\int_\phi d\phi\, \tilde{P} + d\tilde{P}}. \tag{A3}$$

Taking out a factor of $\int_\phi d\phi\, \tilde{P}(\phi)$ in the denominator leads to

$$P(\phi) + dP(\phi) = \frac{\tilde{P} + d\tilde{P}}{\int_\phi d\phi\, \tilde{P}\left[1 + (1/\int_\phi d\phi\, \tilde{P})\int_\phi d\phi\, d\tilde{P}\right]}. \tag{A4}$$

Expanding the expression in the denominator within the square brackets as a power series using the binomial theorem $[(1+x)^n = 1 + nx + n(n-1)x^2/2 + \cdots]$, yields

$$P(\phi) + dP(\phi) \simeq \frac{\tilde{P} + d\tilde{P}}{\int_\phi d\phi\, \tilde{P}}\left[1 - \frac{\int_\phi d\phi\, d\tilde{P}}{\int_\phi d\phi\, \tilde{P}} + \frac{\left(\int_\phi d\phi\, d\tilde{P}\right)^2}{\left(\int_\phi d\phi\, \tilde{P}\right)^2}\right]. \tag{A5}$$

Normalizing the distribution $\tilde{P}$ using the factors of $\int_\phi d\phi\, P(\phi)$ in the denominator and also substituting in the expression for dP in Eq. (A1), we obtain

$$P + dP = [P + (XI + \text{c.c.})P\, dt]\left[1 - \int_\phi d\phi (XI + \text{c.c.})P(\phi)dt + \left(\int_\phi d\phi (XI + \text{c.c.})P(\phi)dt\right)^2\right] \tag{A6}$$

$$= [P + (XI + \text{c.c.})P\, dt][1 - (\langle X\rangle_P I + \text{c.c.})dt + (\langle X\rangle_P I + \text{c.c.})^2 dt^2]. \tag{A7}$$

Expanding this expression and keeping only terms of order dt or less, we arrive at Eq. (A2).

APPENDIX B

In this appendix we demonstrate that, for the schemes based on quantum parameter estimation (the optimal heterodyne-based, the canonical, and the semioptimal adaptive schemes),

$$\langle e^{i(\phi - \hat{\phi})}\rangle_{\xi,I} = \langle |\langle e^{i\phi}\rangle_{P(\phi)}|\rangle_I. \tag{B1}$$

By definition

$$\langle e^{i(\phi - \hat{\phi})}\rangle_{\xi,I} = \int_\xi \int_I d\xi\, dI\, P(\xi,I)e^{i[\phi(\xi) - \hat{\phi}(I)]}. \tag{B2}$$

Expressing $e^{i[\phi(\xi) - \hat{\phi}(I)]}$ as an integral over the dummy phase variable φ, we obtain

$$e^{i[\phi(\xi) - \hat{\phi}(I)]} = \int_\varphi d\varphi\, \delta(\phi(\xi) - \varphi)e^{i[\varphi - \hat{\phi}(I)]}. \tag{B3}$$

Substituting the right-hand side of Eq. (B3) into the right-hand side of Eq. (B2) yields

$$\langle e^{i(\phi - \hat{\phi})}\rangle_{\xi,I} = \int_\xi \int_I \int_\varphi d\xi\, dI\, d\varphi\, P(\xi)P(I|\xi)\delta(\phi(\xi) - \varphi)e^{i[\varphi - \hat{\phi}(I)]}. \tag{B4}$$

Assuming we know the so-called process noise ξ, then we know the phase ϕ exactly and thus our probability density function for ϕ is a Dirac δ function. From this it follows that

$$P(I|\xi)\delta(\phi(\xi) - \varphi)d\varphi = P(\varphi, I|\xi)d\varphi. \tag{B5}$$

Substituting this result into Eq. (B4) and integrating over ξ yields

$$\langle e^{i(\phi-\hat{\phi})}\rangle_{\xi,I} = \int_I \int_\varphi dI\, d\varphi\, P(\varphi,I) e^{i[\varphi-\hat{\phi}(I)]}. \tag{B6}$$

Using elementary probability theory, we obtain

$$\langle e^{i(\phi-\hat{\phi})}\rangle_{\xi,I} = \int_I dI\, P(I) \int_\varphi d\varphi\, P(\varphi|I) e^{i[\varphi-\hat{\phi}(I)]}. \tag{B7}$$

Given that

$$\hat{\phi}(I) = \arg\left(\int d\varphi'\, P(\varphi'|I) e^{i\varphi'} \right), \tag{B8}$$

where φ' is a second dummy phase variable, Eq. (B7) leads to

$$\langle e^{i(\phi-\hat{\phi})}\rangle_{\xi,I} = \int_I dI\, P(I) \left| \int_\varphi d\varphi\, P(\varphi|I) e^{i\varphi} \right| = \langle |\langle e^{i\varphi}\rangle_{P(\varphi)}|\rangle_I. \tag{B9}$$

Upon replacing φ by ϕ in the final expression, where ϕ now acts as a dummy phase variable, Eq. (B1) is obtained.

[1] M. J. Hall and I. G. Fuss, Quantum Opt. **3**, 147 (1991).
[2] C. M. Caves and P. D. Drummond, Rev. Mod. Phys. **66**, 481 (1994).
[3] H. Mabuchi (private communication).
[4] S. L. Braunstein, A. S. Lane, and C. M. Caves, Phys. Rev. Lett. **69**, 2153 (1992); B. C. Sanders and G. J. Milburn, *ibid.* **75**, 2944 (1995); H. Mabuchi, Quantum Semiclassic. Opt. **8**, 1103 (1996); F. Verstraete, A. C. Doherty, and H. Mabuchi, Phys. Rev. A **64**, 032111 (2001); P. Warszawski, J. Gambetta, and H. M. Wiseman, *ibid.* **69**, 042104 (2004).
[5] H. W. Wiseman, Phys. Rev. Lett. **75**, 4587 (1995).
[6] H. M. Wiseman and R. B. Killip, Phys. Rev. A **56**, 944 (1997).
[7] H. M. Wiseman and R. B. Killip, Phys. Rev. A **57**, 2169 (1998).
[8] D. Berry and H. M. Wiseman, Phys. Rev. A **65**, 043803 (2002).
[9] M. A. Armen *et al.*, Phys. Rev. Lett. **89**, 133602 (2002).
[10] H. P. Yuen and J. H. Shapiro, IEEE Trans. Inf. Theory **IT-24**, 675 (1978).
[11] H. P. Yuen and J. H. Shapiro, IEEE Trans. Inf. Theory **IT-26**, 78 (1978).
[12] J. H. Shapiro, H. P. Yuen, and J. A. Machado Mata, IEEE Trans. Inf. Theory **IT-25**, 179 (1979).
[13] J. H. Shapiro and S. S. Wagner, IEEE J. Quantum Electron. **QE-20**, 803 (1984).
[14] J. H. Shapiro, IEEE J. Quantum Electron. **QE-21**, 237 (1985).
[15] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer, New York, 1983), Chap. 4.
[16] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, U.K., 1995).
[17] U. Leonhardt, J. A. Vaccaro, B. Böhmer, and H. Paul, Phys. Rev. A **51**, 84 (1995).
[18] S.M. Barnett and B. J. Dalton, Phys. Scr., T **48**, 13 (1993).
[19] D. T. Pegg and S. M. Barnett, Europhys. Lett. **6**, 483 (1988).
[20] A. Yariv, *Quantum Electronics* (John Wiley & Sons, New York, 1989).
[21] A. Benssousan, *Stochastic Control of Partially Observable Systems* (Cambridge University Press, Cambridge, U.K., 1992).
[22] T. P. McGarty, *Stochastic Systems and State Estimation* (Addison-Wesley, Sydney, 1974).
[23] A. S. Holevo, Lect. Notes Math. **1055**, 153 (1994).
[24] A. Doherty *et al.*, Phys. Rev. A **62**, 012105 (2000).
[25] H. M. Wiseman, Phys. Rev. A **47**, 5180 (1993).
[26] D. W. Berry and H. M. Wiseman, J. Mod. Opt. **48**, 797 (2001).
[27] D. T. Pope, D. W. Berry, and N. K. Langford (unpublished).
[28] J. K. Stockton, J. M. Gerimia, A. C. Doherty, and H. Mabuchi, Phys. Rev. A **69**, 032109 (2004).