

Quantum Bell-Ziv-Zakai Bounds and Heisenberg Limits for Waveform Estimation

Dominic W. Berry,¹ Mankei Tsang,^{2,3} Michael J. W. Hall,⁴ and Howard M. Wiseman⁴

¹*Department of Physics and Astronomy, Macquarie University, Sydney NSW 2109, Australia*

²*Department of Electrical and Computer Engineering, National University of Singapore,*

4 Engineering Drive 3, Singapore 117583, Singapore

³*Department of Physics, National University of Singapore,*

2 Science Drive 3, Singapore 117551, Singapore

⁴*Centre for Quantum Computation and Communication Technology (Australian Research Council),*

Centre for Quantum Dynamics, Griffith University, Brisbane QLD 4111, Australia

(Received 14 October 2014; revised manuscript received 18 May 2015; published 18 August 2015)

We propose quantum versions of the Bell-Ziv-Zakai lower bounds for the error in multiparameter estimation. As an application, we consider measurement of a time-varying optical phase signal with stationary Gaussian prior statistics and a power-law spectrum $\sim 1/|\omega|^p$, with $p > 1$. With no other assumptions, we show that the mean-square error has a lower bound scaling as $1/\mathcal{N}^{2(p-1)/(p+1)}$, where \mathcal{N} is the time-averaged mean photon flux. Moreover, we show that this scaling is achievable by sampling and interpolation, for any $p > 1$. This bound is thus a rigorous generalization of the Heisenberg limit, for measurement of a single unknown optical phase, to a stochastically varying optical phase.

DOI: [10.1103/PhysRevX.5.031018](https://doi.org/10.1103/PhysRevX.5.031018)

Subject Areas: Photonics, Quantum Physics

I. INTRODUCTION

The probabilistic nature of quantum mechanics imposes fundamental limits to hypothesis testing and parameter estimation [1–5]. Such limits are relevant to many metrological applications, such as optical interferometry, optomechanical sensing, gravitational-wave detection [6–9], optical imaging [10–12], magnetometry, gyroscopy, and atomic clocks [13]. The ultimate quantum limits to parameter estimation have been studied extensively in recent years, as they imply that a minimum amount of resource, such as the average photon number for optical phase estimation, is needed to achieve a desired precision, regardless of the measurement method.

For the measurement of a single optical phase parameter, the ultimate quantum limit to the mean-square error scales as $1/\bar{n}^2$, where \bar{n} is the average photon number of the field which undergoes that phase shift. This scaling is often called the Heisenberg limit. After years of speculation and debate [13–22], the Heisenberg limit for single-parameter linear phase estimation has only recently been proven [23–29]. Although decoherence, such as optical loss and dephasing, can impose stricter limitations [9,30–36], the Heisenberg limit is a more fundamental bound and will be increasingly relevant as quantum technologies continue to improve and decoherence effects are further reduced.

Many real-world tasks, such as optical imaging [37,38], quantum tomography and system identification [39,40], and time-varying signal (waveform) estimation [7,9,41–44], require the estimation of multiple parameters. Multiparameter quantum Cramér-Rao bounds have been known since the 1970s [7,45–47], but efforts to derive multiparameter Heisenberg limits from these bounds have not been successful. This is not surprising since, even in the case of single-parameter phase estimation, it is not possible to derive the Heisenberg limit from the quantum Cramér-Rao bound without additional assumptions on the state [48].

In Ref. [41], some of us recently proposed a Heisenberg-style limit for the estimation of an optical phase waveform with stationary Gaussian prior statistics and a power-law spectrum. However, that limit, being derived from a quantum Cramér-Rao bound, requires additional assumptions: It applies only to the specific class of optical beams described by Gaussian fields, with statistics that are both stationary and time symmetric. A very different approach is given in Ref. [49], which derives a multiparameter Heisenberg limit for independent parameters by applying the single-parameter Heisenberg limit to each parameter. In practice, multiple parameters often have nontrivial prior correlations, particularly in the case of continuous waveform estimation, where the correlations are crucial to pose the problem [50]. Thus, the existence of general Heisenberg limits for such cases has remained an open question.

In this paper, we derive new quantum bounds on multiparameter estimation by developing quantum versions of the Bell-Ziv-Zakai bounds [51]. The Bell-Ziv-Zakai bounds were proposed in 1997 by Bell *et al.* [52], building upon the Ziv-Zakai bound [53], and further generalized by Basu

Published by the American Physical Society under the terms of the *Creative Commons Attribution 3.0 License*. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

and Bresler [54]. We then apply our bounds to the notable task of quantum optical phase waveform estimation. Here, the waveform to be estimated is a time-varying phase shift signal $X(t)$ applied to an optical beam. For a waveform $X(t)$ with stationary Gaussian prior statistics and a power-law spectrum ($\propto 1/|\omega|^p$, $p > 1$), we derive a lower bound on the mean-square error with a $1/\mathcal{N}^{2(p-1)/(p+1)}$ scaling, where \mathcal{N} is the mean photon flux. This proof confirms that the scaling previously proposed in Ref. [41] is valid for arbitrary quantum states. Moreover, we show that this scaling is achievable for all $p > 1$. Previously, achievability has been shown only numerically, and only for $p = 2$ [55]. By contrast, the results in the current paper are rigorous Heisenberg bounds, which are both applicable to arbitrary field states and achievable up to a constant prefactor for all $p > 1$.

This paper is separated into two main parts. The first part, Secs. II–V, assumes unbounded parameters and focuses on the mean-square error. The second part, Secs. VI–VII, focuses on periodic figures of merit, which are more appropriate for periodic parameters such as phase or orientation angles for gyroscopy. They are also insensitive to phase-wrap errors and enable us to demonstrate that our bounds are achievable in more general situations.

II. QUANTUM BELL-ZIV-ZAKAI BOUNDS

A. Classical estimation

First, we summarize known results for the classical estimation problem, following Refs. [52,56], and then present new quantum versions of the bounds in Sec. II B. Let \mathbf{X} be a column vector of unknown real parameters given by

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_J \end{pmatrix}, \quad (2.1)$$

and let \mathbf{Y} be a column vector of observations given by

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_K \end{pmatrix}. \quad (2.2)$$

Both \mathbf{X} and \mathbf{Y} are random variables; we denote the prior probability density by $P_{\mathbf{X}}(\mathbf{x})$ and the conditional observation probability density by $P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$. In other words, $P_{\mathbf{X}}(\mathbf{x})$ (for example) denotes the probability *density* that $\mathbf{X} = \mathbf{x}$, as distinct from $\Pr(e)$, which we use to denote the probability of an event e . Furthermore, let $\check{\mathbf{X}}(\mathbf{Y})$ be an estimator of the unknown \mathbf{X} from the observations \mathbf{Y} . We use the vee accent ($\check{\mathbf{X}}$) for the estimator, rather than the hat ($\hat{\mathbf{X}}$) common in statistics, to avoid confusion with quantum operators.

Define the error vector for the estimator as

$$\boldsymbol{\epsilon}(\mathbf{X}, \mathbf{Y}) := \check{\mathbf{X}}(\mathbf{Y}) - \mathbf{X}. \quad (2.3)$$

To characterize the performance of the estimator, we consider an error measure, or “distortion function,” of the form $D(\mathbf{u}^\top \boldsymbol{\epsilon})$, where \mathbf{u} is an arbitrary real vector with J elements that defines and weights the error components of interest. Here, \top denotes the transpose, so $\mathbf{u}^\top \boldsymbol{\epsilon} = \sum_j u_j \epsilon_j$. For example, the choices $D(\tau) = \tau^2$ and $u_j = \delta_{jk}$ yield the distortion function

$$D_k = [\check{X}_k(\mathbf{Y}) - X_k]^2. \quad (2.4)$$

The expectation value $\mathbb{E}[D_k]$ in this case is the mean-square error in the estimate for the parameter X_k . The notation \mathbb{E} denotes expectation over the random variables \mathbf{X} and \mathbf{Y} . More generally, the choice $D(\tau) = \tau^2$, with arbitrary \mathbf{u} , yields the mean-square error

$$\mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] = \mathbb{E}[(\mathbf{u}^\top \boldsymbol{\epsilon})^2] = \mathbf{u}^\top \boldsymbol{\Sigma} \mathbf{u}, \quad (2.5)$$

where $\boldsymbol{\Sigma}$ is the error covariance matrix

$$\boldsymbol{\Sigma} := \mathbb{E}(\boldsymbol{\epsilon} \boldsymbol{\epsilon}^\top). \quad (2.6)$$

We focus on distortion functions with the properties that they are: nondecreasing on $[0, \infty)$; differentiable, with a derivative denoted by $\dot{D}(z)$; symmetric, that is, $D(z) = D(-z)$; and satisfying $D(0) = 0$. We can then express the expected distortion with respect to the probability density $P_{|\mathbf{u}^\top \boldsymbol{\epsilon}|}(z)$, for the random variable $|\mathbf{u}^\top \boldsymbol{\epsilon}|$, as

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] &= \mathbb{E}[D(|\mathbf{u}^\top \boldsymbol{\epsilon}|)] \\ &= \int_0^\infty dz D(z) P_{|\mathbf{u}^\top \boldsymbol{\epsilon}|}(z). \end{aligned} \quad (2.7)$$

Since the probability density is equal to the negative derivative of the cumulative probability function $\Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq z)$, we obtain, via integration by parts,

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] &= [-D(z) \Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq z)] \Big|_0^\infty \\ &\quad + \int_0^\infty dz \dot{D}(z) \Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq z) \\ &= \frac{1}{2} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \Pr\left(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \frac{\tau}{2}\right). \end{aligned} \quad (2.8)$$

In the last line, we have made the substitution $z = \tau/2$, to conform with the notation of Ref. [52].

Since \dot{D} is nonnegative by assumption, a lower bound on the expected distortion can be obtained by lower bounding the probability $\Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2)$. The central idea of the Ziv-Zakai family of bounds is to bound $\Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2)$ using the error probability in a related binary hypothesis-testing

problem. To this end, let \mathbf{v} be a real column vector with J elements that satisfies the restriction

$$\mathbf{u}^\top \mathbf{v} = 1 \quad (2.9)$$

with respect to the error-weighting vector \mathbf{u} . This is an important restriction on \mathbf{v} which we will use throughout the paper. We rewrite $\Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2)$ as follows:

$$\begin{aligned} & \Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2) \\ &= \Pr(\mathbf{u}^\top \boldsymbol{\epsilon} > \tau/2) + \Pr(\mathbf{u}^\top \boldsymbol{\epsilon} \leq -\tau/2) \\ &= \int d\mathbf{x} P_X(\mathbf{x}) \Pr(\mathbf{u}^\top \boldsymbol{\epsilon} > \tau/2 | \mathbf{X} = \mathbf{x}) \\ &\quad + \int d\mathbf{x} P_X(\mathbf{x}) \Pr(\mathbf{u}^\top \boldsymbol{\epsilon} \leq -\tau/2 | \mathbf{X} = \mathbf{x}) \\ &= \int d\mathbf{x} P_X(\mathbf{x}) \Pr(\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) > \mathbf{u}^\top \mathbf{x} + \tau/2 | \mathbf{X} = \mathbf{x}) \\ &\quad + \int d\mathbf{x} P_X(\mathbf{x} + \mathbf{v}\tau) \\ &\quad \times \Pr(\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) \leq \mathbf{u}^\top (\mathbf{x} + \mathbf{v}\tau) - \tau/2 | \mathbf{X} = \mathbf{x} + \mathbf{v}\tau) \quad (2.10) \end{aligned}$$

$$\begin{aligned} &= \int d\mathbf{x} [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] \\ &\quad \times [\pi_0 \Pr(\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) > \mathbf{u}^\top \mathbf{x} + \tau/2 | \mathbf{X} = \mathbf{x}) \\ &\quad + \pi_1 \Pr(\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) \leq \mathbf{u}^\top (\mathbf{x} + \mathbf{v}\tau) - \tau/2 | \mathbf{X} = \mathbf{x} + \mathbf{v}\tau)], \quad (2.11) \end{aligned}$$

where

$$\pi_0 := \frac{P_X(\mathbf{x})}{P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)}, \quad (2.12)$$

$$\pi_1 := 1 - \pi_0. \quad (2.13)$$

In Eq. (2.11), we have used the restriction that $\mathbf{u}^\top \mathbf{v} = 1$.

The term in square brackets in the last two lines of Eq. (2.11) is the probability of error for distinguishing \mathbf{x} from $\mathbf{x} + \mathbf{v}\tau$ with prior probabilities π_0 and π_1 , respectively, from the observations \mathbf{Y} and a particular decision rule. This decision rule is that, if $\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) \leq \mathbf{u}^\top \mathbf{x} + \tau/2$, the decision is taken to choose \mathbf{x} , and if $\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) > \mathbf{u}^\top \mathbf{x} + \tau/2$, the decision is taken to choose $\mathbf{x} + \mathbf{v}\tau$. Because this is the probability of error for a given decision rule, it cannot be smaller than the *minimum* probability of error, which we denote $P_e(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau)$. To be explicit, this minimum probability of error is given by [57,58]

$$\begin{aligned} & P_e(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau) \\ &= \frac{1}{2} - \frac{1}{2} \int d\mathbf{y} |\pi_0 P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) - \pi_1 P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x} + \mathbf{v}\tau)|. \quad (2.14) \end{aligned}$$

Therefore, we obtain the lower bound

$$\begin{aligned} & \Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2) \\ & \geq \int d\mathbf{x} [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] P_e(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau). \quad (2.15) \end{aligned}$$

This bound holds for any \mathbf{v} satisfying $\mathbf{u}^\top \mathbf{v} = 1$. Thus, we can also take the supremum over such \mathbf{v} 's to obtain the stronger lower bound,

$$\begin{aligned} & \Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2) \\ & \geq \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v} = 1} \int d\mathbf{x} [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] P_e(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau). \quad (2.16) \end{aligned}$$

Substituting Eq. (2.16) into Eq. (2.8) then gives the Bell-Ziv-Zakai bound [52,56]

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] & \geq \frac{1}{2} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \\ & \quad \times \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v} = 1} \int d\mathbf{x} [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] \\ & \quad \times P_e(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau). \quad (2.17) \end{aligned}$$

An alternative lower bound can be obtained by replacing $P_X(\mathbf{x})$ and $P_X(\mathbf{x} + \mathbf{v}\tau)$ with $\min\{P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)\}$ in Eq. (2.10), which gives

$$\begin{aligned} & \Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2) \\ & \geq 2 \int d\mathbf{x} \min\{P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)\} \\ & \quad \times \left[\frac{1}{2} \Pr(\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) > \mathbf{u}^\top \mathbf{x} + \tau/2 | \mathbf{X} = \mathbf{x}) \right. \\ & \quad \left. + \frac{1}{2} \Pr(\mathbf{u}^\top \check{\mathbf{X}}(\mathbf{Y}) \leq \mathbf{u}^\top (\mathbf{x} + \mathbf{v}\tau) - \tau/2 | \mathbf{X} = \mathbf{x} + \mathbf{v}\tau) \right]. \quad (2.18) \end{aligned}$$

This time, the expression in square brackets is the error probability for a decision problem with *equal* prior probabilities. Let $P_e^{\text{el}}(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau)$ denote the minimum error probability for the decision problem with equal prior probabilities. It can be obtained from Eq. (2.14) by substituting $\pi_0 = \pi_1 = 1/2$. This yields

$$\begin{aligned} & \Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2) \\ & \geq 2 \int d\mathbf{x} \min\{P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)\} P_e^{\text{el}}(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau). \quad (2.19) \end{aligned}$$

Using this result in Eq. (2.8) then gives the alternative Bell-Ziv-Zakai bound [52,56]

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] &\geq \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int d\mathbf{x} \\ &\times \min[P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)] P_e^{\text{el}}(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau). \end{aligned} \quad (2.20)$$

For simplicity, here we have omitted a ‘‘valley-filling’’ function used in Ref. [56] that can, in principle, tighten the bounds further, but makes analytic results more difficult to obtain; see Appendix A for the tighter bounds.

Note that, in both Eqs. (2.15) and (2.19), $\Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2)$ is lower bounded by an overlap integral between two terms. The first term depends only on *prior* information: the prior probability density $P_X(\mathbf{x})$ and its translated form $P_X(\mathbf{x} + \mathbf{v}\tau)$. The second term further depends on *posterior* information: the degree to which the observations \mathbf{Y} can distinguish $\mathbf{X} = \mathbf{x}$ from $\mathbf{X} = \mathbf{x} + \mathbf{v}\tau$. Finding the translation vector \mathbf{v} that maximizes the overlap, subject to the constraint $\mathbf{u}^\top \mathbf{v} = 1$ for a given error-weighting vector \mathbf{u} , will give the tightest lower bounds.

This concludes the summary of prior results for the classical case, and we next present our bounds for quantum estimation.

B. Quantum estimation

For the quantum parameter estimation problem, let $\hat{\rho}_X$ be the density operator that describes the state of a quantum probe as a function of the unknown parameters \mathbf{X} , and let $\hat{E}_Y(\mathbf{y})$ be the positive operator-valued measure (POVM) that specifies the observation statistics [1]. The likelihood function is then

$$P_{Y|X}(\mathbf{y}|\mathbf{x}) = \text{tr}[\hat{E}_Y(\mathbf{y})\hat{\rho}_x], \quad (2.21)$$

with tr denoting the operator trace. It is known [1,28] that, for any POVM,

$$P_e(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau) \geq \frac{1}{2} - \frac{1}{2} \|\pi_0 \hat{\rho}_x - \pi_1 \hat{\rho}_{x+\mathbf{v}\tau}\|_1 \quad (2.22)$$

$$\geq \frac{1}{2} \left[1 - \sqrt{1 - 4\pi_0\pi_1 F(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau)} \right], \quad (2.23)$$

where

$$\|\hat{A}\|_1 := \text{tr} \sqrt{\hat{A}^\dagger \hat{A}} \quad (2.24)$$

is the trace norm and

$$F(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau) := \left(\text{tr} \sqrt{\sqrt{\hat{\rho}_x} \hat{\rho}_{x+\mathbf{v}\tau} \sqrt{\hat{\rho}_x}} \right)^2 \quad (2.25)$$

is the Uhlmann fidelity [59,60]. Equations (2.22) and (2.23), together with the first Bell-Ziv-Zakai bound given by Eq. (2.17), then give quantum lower bounds on the estimation error:

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] &\geq \frac{1}{4} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \\ &\times \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int d\mathbf{x} [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] \\ &\times (1 - \|\pi_0 \hat{\rho}_x - \pi_1 \hat{\rho}_{x+\mathbf{v}\tau}\|_1) \end{aligned} \quad (2.26)$$

$$\begin{aligned} &\geq \frac{1}{4} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \\ &\times \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int d\mathbf{x} [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] \\ &\times [1 - \sqrt{1 - 4\pi_0\pi_1 F(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau)}]. \end{aligned} \quad (2.27)$$

Similarly, Eq. (2.20) and quantum bounds on P_e^{el} via Eqs. (2.22) and (2.23) (with $\pi_0 = \pi_1 = 1/2$) lead to the lower bounds

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] &\geq \frac{1}{2} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \\ &\times \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int d\mathbf{x} \min[P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)] \\ &\times \left(1 - \frac{1}{2} \|\hat{\rho}_x - \hat{\rho}_{x+\mathbf{v}\tau}\|_1 \right) \end{aligned} \quad (2.28)$$

$$\begin{aligned} &\geq \frac{1}{2} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \\ &\times \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int d\mathbf{x} \min[P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)] \\ &\times [1 - \sqrt{1 - F(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau)}]. \end{aligned} \quad (2.29)$$

We call Eqs. (2.26)–(2.29) the quantum Bell-Ziv-Zakai bounds.

C. Special conditions: Homogeneous displacements and Gaussian prior

To derive further analytic results, we focus on the case in which the quantum probe state $\hat{\rho}_x$ is related to an input probe state $\hat{\rho}$ by

$$\hat{\rho}_x = \alpha_x(\hat{\rho}), \quad (2.30)$$

where $\{\alpha_x\}$ is a set of completely positive trace-preserving (CPTP) maps satisfying the homogeneity condition

$$\alpha_x \circ \alpha_{x'} = \alpha_{x+x'}. \quad (2.31)$$

Here, the addition can be normal vector addition, or it can be modulo addition. This includes, for example, unitary transformations such as optical phase shifts and phase-space translations (see Secs. III and IV below), as well as homogeneous decoherence processes such as the addition of Gaussian noise to optical modes (with \mathbf{x} being the noise

variances) [61] or lossy transmission of optical modes (with \mathbf{x} being a logarithmic measure of transmission).

Now, since the trace norm cannot increase under CPTP maps [60], one has $\|\pi_0 \hat{\rho}_{\mathbf{x}} - \pi_1 \hat{\rho}_{\mathbf{x}+\mathbf{v}\tau}\|_1 = \|\alpha_{\mathbf{x}}(\pi_0 \hat{\rho} - \pi_1 \hat{\rho}_{\mathbf{v}\tau})\|_1 \leq \|\pi_0 \hat{\rho} - \pi_1 \hat{\rho}_{\mathbf{v}\tau}\|_1$, implying that $\hat{\rho}_{\mathbf{x}}$ and $\hat{\rho}_{\mathbf{x}+\mathbf{v}\tau}$ may be replaced by $\hat{\rho}$ and $\hat{\rho}_{\mathbf{v}\tau}$, respectively, in each of the quantum Ziv-Zakai bounds in Eqs. (2.26) and (2.28). Similarly, since the fidelity is nondecreasing under such maps [60], one may replace $F(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau)$ by $F(0, \mathbf{v}\tau)$ in the bounds (2.27) and (2.29). As a consequence, the integration over \mathbf{x} in these bounds greatly simplifies, being dependent only on the prior probability densities $P_{\mathbf{X}}(\mathbf{x})$ and $P_{\mathbf{X}}(\mathbf{x} + \mathbf{v}\tau)$.

Next, we impose the condition that $P_{\mathbf{X}}(\mathbf{x})$ is a multivariate Gaussian distribution. In terms of its covariance matrix

$$\Sigma_0 := \mathbb{E}(\mathbf{X}\mathbf{X}^\top) - \mathbb{E}(\mathbf{X})\mathbb{E}(\mathbf{X}^\top), \quad (2.32)$$

the integral over \mathbf{x} in Eqs. (2.28) and (2.29) simplifies to [52]

$$\int d\mathbf{x} \min[P_{\mathbf{X}}(\mathbf{x}), P_{\mathbf{X}}(\mathbf{x} + \mathbf{v}\tau)] = \text{erfc}\left(\frac{\sqrt{\pi}\tau}{2\tau_0(\mathbf{v})}\right), \quad (2.33)$$

where

$$\text{erfc}z := \frac{2}{\sqrt{\pi}} \int_z^\infty d\zeta \exp(-\zeta^2), \quad (2.34)$$

$$\tau_0(\mathbf{v}) := \sqrt{\frac{2\pi}{\mathbf{v}^\top \Sigma_0^{-1} \mathbf{v}}} \quad (2.35)$$

(see also Appendix B). Note from Eq. (2.33) that $\tau_0(\mathbf{v})$ sets the approximate scale of τ above for which the prior density $P_{\mathbf{X}}(\mathbf{x})$ can be reliably distinguished from the translated density $P_{\mathbf{X}}(\mathbf{x} + \mathbf{v}\tau)$.

A convenient lower bound on the complementary error function, useful for obtaining analytic results in the next section, is

$$\text{erfc}\left(\frac{\sqrt{\pi}\tau}{2\tau_0(\mathbf{v})}\right) \geq \Lambda\left(\frac{\tau}{\tau_0(\mathbf{v})}\right), \quad (2.36)$$

where Λ is the triangle function $\Lambda(z) := \max(1 - |z|, 0)$, as shown in Fig. 1. Thus, for homogeneous displacements having a Gaussian prior probability density, the quantum Ziv-Zakai bounds in Eqs. (2.28) and (2.29) may be replaced by the simplified bounds

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] &\geq \frac{1}{2} \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \Lambda\left(\frac{\tau}{\tau_0(\mathbf{v})}\right) \\ &\times \left(1 - \frac{1}{2} \|\hat{\rho} - \hat{\rho}_{\mathbf{v}\tau}\|_1\right) \end{aligned} \quad (2.37)$$

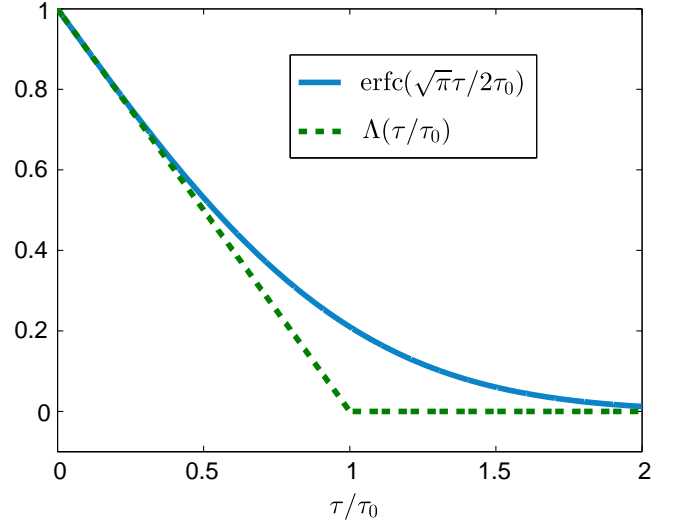


FIG. 1. The erfc function and a lower bound using the triangle function Λ .

$$\begin{aligned} &\geq \frac{1}{2} \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \Lambda\left(\frac{\tau}{\tau_0(\mathbf{v})}\right) \\ &\times \left(1 - \sqrt{1 - F(0, \mathbf{v}\tau)}\right). \end{aligned} \quad (2.38)$$

III. ESTIMATION FOR COMMUTING GENERATORS WITH GAUSSIAN PRIOR

We now consider the problem of probe states displaced by commuting generators, as illustrated in Fig. 2, as well as a Gaussian prior distribution for \mathbf{X} . For this problem, the output probe state has the general form

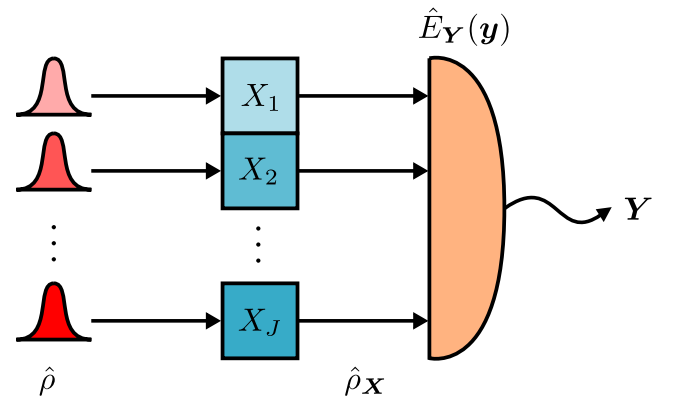


FIG. 2. The quantum estimation problem for commuting generators. The initial state ρ may be entangled, and each subsystem or mode undergoes a displacement X_j generated by operator \hat{g}_j , with $[\hat{g}_j, \hat{g}_k] = 0$. The output may be measured via some general joint measurement $\hat{E}_Y(\mathbf{y})$.

$$\hat{\rho}_X = \alpha_X(\hat{\rho}) := \exp(-iX^\top \hat{\mathbf{g}})\hat{\rho}\exp(iX^\top \hat{\mathbf{g}}), \quad (3.1)$$

where $\hat{\rho}$ is the initial quantum state and $\hat{\mathbf{g}}$ is a column vector of shift generators with $[\hat{g}_j, \hat{g}_k] = 0$. The operators \hat{g}_j must be Hermitian but are otherwise unrestricted. Note this includes multimode phase estimation as a special case, with \hat{g}_j corresponding to the photon number operator \hat{n}_j of mode j , and X_j to the phase shift generated by \hat{n}_j .

Equation (3.1) satisfies the homogeneous displacement property in Eq. (2.31), and more generally, it would satisfy this property for $[\hat{g}_j, \hat{g}_k] = c_{jk} = \text{constant}$. With the assumption of a Gaussian prior, the simplified quantum Ziv-Zakai bounds in Eqs. (2.37) and (2.38) are applicable. We now simplify the latter of these bounds further, by relating the fidelity term $F(0, \mathbf{v}\tau)$ to the expectation value of $\hat{\mathbf{g}}$.

First, note that if $|\psi\rangle$ is a purification of $\hat{\rho}$ on some larger Hilbert space, then $\exp(i\tau\mathbf{v}^\top \hat{\mathbf{g}})|\psi\rangle$ is a purification of $\hat{\rho}_{\mathbf{v}\tau}$. Hence, defining $\langle \hat{O} \rangle := \text{tr}(\hat{O}\hat{\rho})$ and $C(\mathbf{g}) := \langle \mathbf{g}|\hat{\rho}|\mathbf{g} \rangle$ for eigenstate $|\mathbf{g}\rangle$ of $\hat{\mathbf{g}}$, Uhlmann's theorem [59,60] yields the lower bound

$$\begin{aligned} F(0, \mathbf{v}\tau) &\geq |\langle \exp(-i\tau\mathbf{v}^\top \hat{\mathbf{g}}) \rangle|^2 \\ &= \sum_{\mathbf{g}, \mathbf{g}'} C(\mathbf{g})C(\mathbf{g}') \cos[\tau\mathbf{v}^\top (\mathbf{g} - \mathbf{g}')]. \end{aligned} \quad (3.2)$$

We are ultimately interested in obtaining a Heisenberg limit, i.e., a limit scaling linearly with the modulus of the shift generator. In order to achieve this, we use the linear bound [23]

$$\cos \theta \geq 1 - \lambda|\theta| \quad (3.3)$$

on the cosine function, where $\lambda \approx 0.7246$ is a solution of $\lambda(\pi - \arcsin \lambda) = 1 + \sqrt{1 - \lambda^2}$, as depicted in Fig. 3. This leads to

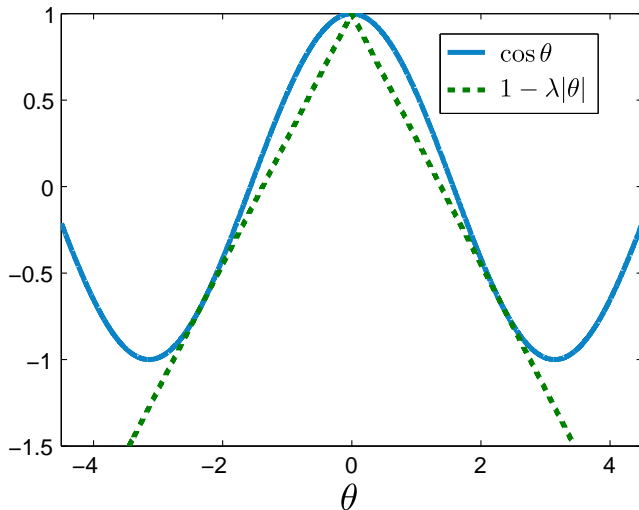


FIG. 3. A lower bound on the cosine function.

$$\begin{aligned} F(0, \mathbf{v}\tau) &\geq \sum_{\mathbf{g}, \mathbf{g}'} C(\mathbf{g})C(\mathbf{g}') [1 - \lambda\tau|\mathbf{v}^\top (\mathbf{g} - \mathbf{g}')|] \\ &\geq \sum_{\mathbf{g}, \mathbf{g}'} C(\mathbf{g})C(\mathbf{g}') [1 - \lambda\tau(|\mathbf{v}^\top \mathbf{g}| + |\mathbf{v}^\top \mathbf{g}'|)] \\ &\geq \sum_{\mathbf{g}, \mathbf{g}'} C(\mathbf{g})C(\mathbf{g}') [1 - \lambda\tau(|\mathbf{v}|^\top |\mathbf{g}| + |\mathbf{v}|^\top |\mathbf{g}'|)] \\ &= 1 - 2\lambda\tau|\mathbf{v}|^\top \langle |\hat{\mathbf{g}}| \rangle, \end{aligned} \quad (3.4)$$

where $|\mathbf{v}|$ is the vector formed from \mathbf{v} by taking the absolute values of each of its elements, and similarly, $|\hat{\mathbf{g}}|$ is the vector formed by replacing each eigenvalue of each element \hat{g}_j of $\hat{\mathbf{g}}$ by its absolute value. Since fidelity is non-negative, a tighter bound is

$$F(0, \mathbf{v}\tau) \geq \Lambda\left(\frac{\tau}{\tau_F(\mathbf{v})}\right) \quad (3.5)$$

in terms of the triangle function defined below Eq. (2.36), where

$$\tau_F(\mathbf{v}) := \frac{1}{2\lambda|\mathbf{v}|^\top \langle |\hat{\mathbf{g}}| \rangle}. \quad (3.6)$$

In analogy to $\tau_0(\mathbf{v})$ in Eq. (2.35), $\tau_F(\mathbf{v})$ sets the approximate scale for τ above which the probe states $\hat{\rho}_x$ and $\hat{\rho}_{x+\mathbf{v}\tau}$ can be reliably distinguished, as measured by the fidelity. Note that a slightly tighter bound, with $|\hat{\mathbf{g}}|$ replaced by $|\hat{\mathbf{g}} - \mathbf{h}|$ for any constant vector \mathbf{h} , may be obtained by replacing $\mathbf{g} - \mathbf{g}'$ by $(\mathbf{g} - \mathbf{h}) - (\mathbf{g}' - \mathbf{h})$ in the first line of Eq. (3.4).

Finally, choosing the distortion function to be $D(\tau) = \tau^2$, so that the expected distortion is given by the mean-square error in Eq. (2.5), we apply Eq. (3.5) in Eq. (2.38) to derive the lower bound

$$\begin{aligned} \mathbf{u}^\top \Sigma \mathbf{u} &\geq \frac{1}{2} \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int_0^\infty d\tau \tau \Lambda\left(\frac{\tau}{\tau_0(\mathbf{v})}\right) \Lambda\left(\sqrt{\frac{\tau}{\tau_F(\mathbf{v})}}\right) \\ &= \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} Z(\mathbf{v}) \end{aligned} \quad (3.7)$$

where

$$Z(\mathbf{v}) := \begin{cases} \tau_F(\mathbf{v})^2 \left(\frac{1}{20} - \frac{\tau_F(\mathbf{v})}{42\tau_0(\mathbf{v})} \right) & \tau_F(\mathbf{v}) \leq \tau_0(\mathbf{v}) \\ \tau_0(\mathbf{v})^2 \left(\frac{1}{12} - \frac{2}{35} \sqrt{\frac{\tau_0(\mathbf{v})}{\tau_F(\mathbf{v})}} \right) & \tau_F(\mathbf{v}) > \tau_0(\mathbf{v}). \end{cases} \quad (3.8)$$

Recall that both τ_0 in Eq. (2.35) and τ_F in Eq. (3.6) vary inversely with the magnitude of \mathbf{v} , ensuring that both sides of Eq. (3.7) vary as the square of the magnitude of \mathbf{u} . While τ_0 relates to the prior information (the narrower the prior distribution, the smaller τ_0 is), τ_F relates to the average magnitude of the generators with respect to the input state (the greater this average is, the smaller τ_F is).

These properties are quite intuitive—the size of the lower bound on the mean-square error decreases with more prior information or for a quantum state with greater potential sensitivity to any given displacement value x .

IV. WAVEFORM PHASE ESTIMATION

A. Taking the continuous limit

We now consider an optical beam where $X(t)$ is a phase signal that is a continuous function of time t , as illustrated in Fig. 4. We use $\Sigma_0(t, t')$ to denote the covariance function for the prior correlations between $X(t)$ and $X(t')$, and we again take the probability distribution to be Gaussian. The requirement that almost all (in a measure sense) realizations $x(t)$ of $X(t)$ be continuous can be ensured by requiring $\Sigma_0(t, t')$ to be continuous at $t = t'$.

We seek a continuous analogue to the quantum Bell-Ziv-Zakai bound obtained above. In other words, we would like to replace all sums in the preceding section by integrals. Thus, the measurement uncertainty, the quantity on which we wish to place a lower bound, is quantified by

$$\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' u(t) \Sigma(t, t') u(t'), \quad (4.1)$$

where $\Sigma(t, t')$ is the error covariance function for the estimates at times t and t' , and $u(t)$ is a given real function. This expression is the continuous analog of $\mathbf{u}^\top \Sigma \mathbf{u}$ in the discrete Bell-Ziv-Zakai bounds above. Similarly, for a continuous phase signal, the sum $\sum_j X_j \hat{n}_j$ for a multimode phase shift in Eq. (3.1) (where we have specialized to the case where \hat{g}_j is the photon number operator \hat{n}_j of mode j) is replaced with an integral using the photon flux operator. In other words, the probe state $\hat{\rho}$ is modulated by the phase signal $X(t)$ to give

$$\hat{\rho}_X = \exp\left(-i \int dt X(t) \hat{I}(t)\right) \hat{\rho} \exp\left(i \int dt X(t) \hat{I}(t)\right), \quad (4.2)$$

where integration is over the real line. The photon flux operator $\hat{I}(t)$ [3,62,63] can be regarded as the $\Delta t \rightarrow 0$ limit

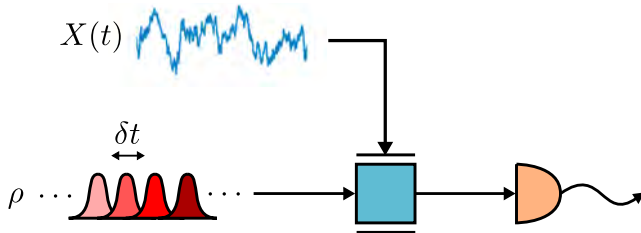


FIG. 4. The continuous waveform estimation problem. The signal $X(t)$ varies continuously, and the system may be approximated by regarding the state as consisting of (typically entangled) modes of length δt , over each of which $X(t)$ does not vary significantly.

of $\Delta \hat{n} / \Delta t$, where $\Delta \hat{n}$ is the number operator corresponding to the photons in a time interval $(t, t + \Delta t)$. Conversely, the integral of $\hat{I}(t)$ over some time interval gives the operator for the number of photons in that time interval.

In the continuum limit, with $\hat{\rho}_X$ given by Eq. (4.2), the bound we wish to derive on the measurement uncertainty as quantified by Eq. (4.1), and the constraints that define that bound, should also be defined in terms of time integrals. But to proceed rigorously, we first need to turn the continuous expressions into discrete expressions by using a small time unit, $\delta t > 0$. This allows us to apply the discrete quantum Bell-Ziv-Zakai bound in Eq. (3.7), prior to taking the limit $\delta t \rightarrow 0$. The general principle is that the state can be approximated by a succession of modes of duration δt such that $X(t)$ does not vary significantly during this time. More specifically, we first approximate $X(t)$ as a constant over time intervals of length δt , so we can use the result (3.7) above and then take the limit as δt approaches zero. Because the realizations of $X(t)$ are continuous, this yields a closer and closer approximation of the true signal as $\delta t \rightarrow 0$.

In particular, consider the discrete times

$$t_j := t_0 + j\delta t, \quad (4.3)$$

where j is an integer and the range of j is over all integers. We can then define corresponding discrete quantities as follows. Approximating $X(t)$ to be constant in each interval $[t_j, t_{j+1})$, we take X_j to be

$$X_j := X(t_j). \quad (4.4)$$

The prior covariance function $\Sigma_0(t, t')$ will then be constant for t and t' within these intervals, and we define

$$\Sigma_{0jk} := \Sigma_0(t_j, t_k). \quad (4.5)$$

We take u_j to be the integral of $u(t)$ over each interval,

$$u_j := \int_{t_j}^{t_{j+1}} dt u(t). \quad (4.6)$$

When $X(t)$ is constant in each interval, it is straightforward to show that the best estimate of $X(t)$ is also constant in each interval. Hence, the error covariance function $\Sigma(t, t')$ can be taken to be constant for t and t' within given time intervals, allowing us to define the discrete covariance matrix

$$\Sigma_{jk} := \Sigma(t_j, t_k). \quad (4.7)$$

It follows that

$$\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' u(t) \Sigma(t, t') u(t') = \mathbf{u}^\top \Sigma \mathbf{u}. \quad (4.8)$$

Finally, we define the operator \hat{g}_j to be

$$\hat{g}_j := \int_{t_j}^{t_{j+1}} \hat{I}(t) dt, \quad (4.9)$$

corresponding to the number of photons between times t_j and t_{j+1} . The integral in Eq. (4.2) then becomes

$$\int_{-\infty}^{\infty} dt X(t) \hat{I}(t) = \mathbf{X}^\top \hat{\mathbf{g}}, \quad (4.10)$$

where now $\hat{\mathbf{g}}$ is the photon number operator $\hat{\mathbf{n}}$. Thus, the integral that generates the shifts of the probe state in Eq. (4.2) is replaced by a sum of the form of the discrete case in Eq. (3.1).

With the above definitions of the discrete quantities, we can use the quantum Bell-Ziv-Zakai bound in Eq. (3.7) to obtain a lower bound on the measurement uncertainty in the continuous case, as quantified by the expression in Eq. (4.1), by taking the limit as $\delta t \rightarrow 0$. However, the bound in Eq. (3.7) depends on $\tau_0(\mathbf{v})$ in Eq. (2.35) and $\tau_F(\mathbf{v})$ in Eq. (3.6), which in turn depend on the quantities $\mathbf{v}^\top \Sigma_0^{-1} \mathbf{v}$ and $|\mathbf{v}|^\top \langle |\hat{\mathbf{g}}| \rangle$, where the vector \mathbf{v} satisfies $\mathbf{u}^\top \mathbf{v} = 1$. Hence, suitable analogues of these quantities are required in the continuous limit.

Therefore, for each value of $\delta t > 0$, we define the corresponding real function $v(t)$ by

$$v(t) := v_j, \quad (4.11)$$

for $t \in [t_j, t_{j+1})$. The constraint $\mathbf{u}^\top \mathbf{v} = 1$ is then equivalent to

$$\int_{-\infty}^{\infty} dt u(t) v(t) = 1. \quad (4.12)$$

Furthermore, noting that the inverse of the prior covariance function, $\Sigma_0^{-1}(t, t')$, is defined via

$$\int_{-\infty}^{\infty} dt' \Sigma_0(t, t') \Sigma_0^{-1}(t', t'') = \delta(t - t''), \quad (4.13)$$

the matrix inverse of Σ_0 in Eq. (4.7) is given by

$$\Sigma_{0jk}^{-1} = \int_{t_j}^{t_{j+1}} dt \int_{t_k}^{t_{k+1}} dt' \Sigma_0^{-1}(t, t'). \quad (4.14)$$

We then obtain

$$\begin{aligned} \frac{2\pi}{\tau_0(\mathbf{v})^2} &= \mathbf{v}^\top \Sigma_0^{-1} \mathbf{v} \\ &= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' v(t) \Sigma_0^{-1}(t, t') v(t') \end{aligned} \quad (4.15)$$

and

$$\frac{1}{2\lambda\tau_F(\mathbf{v})} = |\mathbf{v}|^\top \langle |\hat{\mathbf{g}}| \rangle = \int_{-\infty}^{\infty} dt |v(t)| \langle \hat{I}(t) \rangle, \quad (4.16)$$

where in the latter we have used the fact that $\hat{I}(t)$ is a positive operator for all t . Taking the continuous-time limit $\delta t \rightarrow 0$, the quantum Bell-Ziv-Zakai bound (3.7) thus holds essentially unchanged, with summations replaced by the corresponding integrals in Eqs. (4.1), (4.12), (4.15) and (4.16).

B. Stationarity and power-law scaling

We now make two assumptions, in addition to Gaussianity, about the prior statistics of $X(t)$. The first is stationarity, implying that the prior covariance function $\Sigma_0(t, t')$ is a function of $t - t'$ only. This means that we can define a prior power spectral density $\tilde{\Sigma}_0(\omega)$ by

$$\tilde{\Sigma}_0(\omega) := \int_{-\infty}^{\infty} dt \Sigma_0(t, t') \exp[-i\omega(t - t')]. \quad (4.17)$$

Then, the inverse of $\Sigma_0(t, t')$ is given by

$$\Sigma_0^{-1}(t, t') = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{1}{\tilde{\Sigma}_0(\omega)} \exp[i\omega(t - t')], \quad (4.18)$$

and Eq. (4.15) simplifies to

$$\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' v(t) \Sigma_0^{-1}(t, t') v(t') = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{|\tilde{v}(\omega)|^2}{\tilde{\Sigma}_0(\omega)}, \quad (4.19)$$

with

$$\tilde{v}(\omega) := \int_{-\infty}^{\infty} dt v(t) \exp(-i\omega t). \quad (4.20)$$

The second assumption is that the spectral density $\tilde{\Sigma}_0(\omega)$ has an asymptotic power-law scaling, i.e., $\tilde{\Sigma}_0(\omega) \sim 1/|\omega|^p$ for ω large. This scaling is problematic for small ω , however, because it diverges at $\omega = 0$. To avoid this divergence, it is convenient to take [50]

$$\tilde{\Sigma}_0(\omega) = \frac{\kappa^{p-1}}{|\omega|^p + \gamma^p}, \quad (4.21)$$

for some constant γ . Note that the Ornstein-Uhlenbeck process used in Refs. [42,43] corresponds to $p = 2$. Alternative choices for removing the singularity at $\omega = 0$ yield similar results, as shown in Appendix C. The constant κ essentially gives a multiplicative constant for the phase variation $X(t)$. The power of $p - 1$ ensures that κ has dimensions of reciprocal time. Thus, κ^{-1} is a characteristic time for the spreading; for example, for a

Wiener process, κ^{-1} is the time needed to reach a variance of 1. The constant γ , on the other hand, gives a characteristic time for the relaxation of the phase towards zero and ensures that the spectrum is limited at $\omega = 0$. For example, for an Ornstein-Uhlenbeck process, γ is just the decay constant.

C. Towards the Heisenberg scaling bound

We are particularly interested in the mean-square error of the estimate at a particular time t_0 , i.e., $\Sigma(t_0, t_0)$. For this case, we must choose

$$u(t) = \delta(t - t_0), \quad (4.22)$$

implying from Eq. (4.12) that $v(t_0) = 1$. We can choose any function such that $v(t_0) = 1$ in the quantum Bell-Ziv-Zakai bound. We find it convenient to use the family of functions defined by

$$v(t) = \text{sinc}^2\left(\frac{t - t_0}{2T}\right), \quad (4.23)$$

where

$$\text{sinc } x := \begin{cases} (\sin x)/x & x \neq 0 \\ 1 & x = 0. \end{cases} \quad (4.24)$$

A near-optimal choice for the characteristic time T that parametrizes this family will be made below.

With the above choice of $v(t)$, one has $\tilde{v}(\omega) = 2\pi T e^{i\omega t_0} \Lambda(\omega T)$. This result allows the integral in Eq. (4.19) to be computed analytically for the power spectral density in Eq. (4.21), resulting in this (implicit) expression for τ_0 :

$$\begin{aligned} \frac{2\pi}{\tau_0(\mathbf{v})^2} &= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' v(t) \Sigma_0^{-1}(t, t') v(t') \\ &= \frac{8\pi}{p_3 \kappa^{p-1} T^{p-1}} + \frac{4\pi \gamma^p T}{3\kappa^{p-1}}, \end{aligned} \quad (4.25)$$

where $p_3 := (p+1)(p+2)(p+3)$. We can similarly obtain, from Eq. (4.16), a simple expression for τ_F :

$$\frac{1}{\lambda \tau_F(\mathbf{v})} = \int_{-\infty}^{\infty} dt |v(t)| \langle \hat{I}(t) \rangle = 2\pi T \mathcal{N}(t_0), \quad (4.26)$$

where we have defined a weighted average of the flux around t_0 by

$$\mathcal{N}(t_0) := \frac{\int_{-\infty}^{\infty} dt |v(t)| \langle \hat{I}(t) \rangle}{\int_{-\infty}^{\infty} dt |v(t)|}. \quad (4.27)$$

With these results, we are now in a position to derive a Heisenberg limit.

D. Heisenberg scaling bound: Constant flux

Deriving the desired limit is particularly simple for the case of constant photon flux, $\langle \hat{I}(t) \rangle = \text{constant} = \mathcal{N}$, which also equals $\mathcal{N}(t_0)$. We are interested in the limit of large \mathcal{N} . The lower bound in Eq. (3.7) scales as the lesser of τ_0^2 and τ_F^2 , so the strongest (i.e., largest) lower bound will be obtained for $\tau_0(\mathbf{v}) \approx \tau_F(\mathbf{v})$. Our choice of $v(t)$ in Eq. (4.23) depends on T and t_0 , and the expressions for $\tau_0(\mathbf{v})$ and $\tau_F(\mathbf{v})$ in Eqs. (4.25) and (4.26) only depend on T . Therefore, we wish to choose T to ensure $\tau_0(\mathbf{v}) \approx \tau_F(\mathbf{v})$.

The condition $\tau_0(\mathbf{v}) = \tau_F(\mathbf{v})$ yields

$$\begin{aligned} (2\pi \lambda T \mathcal{N})^2 &= \frac{4}{p_3 \kappa^{p-1} T^{p-1}} + \frac{2\gamma^p T}{3\kappa^{p-1}}, \\ (2\pi \lambda \mathcal{N})^2 &= \frac{4}{p_3 \kappa^{p-1} T^{p+1}} + \frac{2\gamma^p}{3\kappa^{p-1} T}. \end{aligned} \quad (4.28)$$

This equation means that \mathcal{N} is inversely related to T , and the limit of large \mathcal{N} means small T . In that limit, the second term on the right-hand side can be ignored, in which case the solution for the optimal characteristic time T is

$$T = \left(\frac{1}{4\pi^2 \lambda^2 p_3 \kappa^{p-1} \mathcal{N}^2} \right)^{1/(p+1)}. \quad (4.29)$$

We can use this expression for T for all \mathcal{N} , although it is only optimal in the limit of large \mathcal{N} . Then, $\tau_0(\mathbf{v}) < \tau_F(\mathbf{v})$, so using the second subfunction in (3.8) gives

$$Z(\mathbf{v}) \geq \frac{11}{420} \tau_0(\mathbf{v})^2. \quad (4.30)$$

Using this inequality with $\Sigma(t_0, t_0) \geq Z(\mathbf{v})$ gives the lower bound

$$\Sigma(t_0, t_0) \geq c_Z \left(\frac{\kappa}{\mathcal{N}} \right)^{2(p-1)/(p+1)} \left[1 + c_\gamma \left(\frac{\kappa}{\mathcal{N}} \right)^{2p/(p+1)} \right]^{-1}, \quad (4.31)$$

where c_Z and c_γ are the dimensionless constants

$$c_Z := \frac{11}{420} \left(\frac{p_3}{4} \right)^{2/(p+1)} \left(\frac{1}{4\pi\lambda} \right)^{2(p-1)/(p+1)}, \quad (4.32)$$

$$c_\gamma := \frac{p_3}{6} \frac{(\gamma/\kappa)^p}{(4\pi^2 \lambda^2 p_3)^{p/(p+1)}}. \quad (4.33)$$

In the limit of large \mathcal{N} , the expression in square brackets in Eq. (4.31) approaches 1, and we have

$$\Sigma(t_0, t_0) \geq c_Z \left(\frac{\kappa}{\mathcal{N}} \right)^{2(p-1)/(p+1)}. \quad (4.34)$$

Note that this bound on the mean-square error is independent of the constant γ used to remove the singularity at

$\omega = 0$ in Eq. (4.21); i.e., the result depends on the scaling of the spectrum for large ω , not on the behavior for small ω . Alternative choices for removing the singularity yield similar results, as shown in Appendix C.

Note also that both the strict bound in Eq. (4.31) and the asymptotic bound in Eq. (4.34) continue to hold if Eq. (4.21) just gives a lower bound on the spectrum, rather than the exact spectrum. This is because increasing the spectrum will only decrease the value of the integral in Eq. (4.19), leading to a larger value of $\tau_0(\mathbf{v})$. Because $Z(\mathbf{v})$ in Eq. (3.8) is monotonically increasing with $\tau_0(\mathbf{v})$, this increase in $\tau_0(\mathbf{v})$ would only increase the lower bound.

The $(\kappa/\mathcal{N})^{2(p-1)/(p+1)}$ scaling of Eq. (4.34) was previously proposed in Ref. [41] as the Heisenberg limit for a stochastically varying phase with a power-law spectrum. However, the proof in that work applies only to a specific class of Gaussian quantum states. Here, we have proved the scaling for arbitrary quantum states with constant flux by using our powerful new technique of the quantum Bell-Ziv-Zakai bound. Next, we show how to obtain essentially the same result even for a time-varying flux.

E. Heisenberg scaling bound: Varying flux

For the general case of a time-varying photon flux, the overall performance of any estimate is characterized by the error averaged over time, i.e., by

$$\bar{\Sigma} := \lim_{T_{\max} \rightarrow \infty} \frac{1}{2T_{\max}} \int_{-T_{\max}}^{T_{\max}} dt_0 \Sigma(t_0, t_0). \quad (4.35)$$

For constant flux, it is clear that this has the same lower bound as $\Sigma(t_0, t_0)$ in Eq. (4.34) above.

More generally, we define the time-averaged flux by

$$\mathcal{N} := \lim_{T_{\max} \rightarrow \infty} \frac{1}{2T_{\max}} \int_{-T_{\max}}^{T_{\max}} dt_0 \langle \hat{I}(t_0) \rangle. \quad (4.36)$$

It is easy to check that the time average of $\mathcal{N}(t_0)$ in Eq. (4.27) is also equal to \mathcal{N} , implying that this definition is consistent with the above notation for the case of constant flux.

We can rewrite the lower bound in Eq. (3.7) as

$$\Sigma(t_0, t_0) \geq \tau_0(\mathbf{v})^2 \times \begin{cases} \left[\frac{\tau_F(\mathbf{v})}{\tau_0(\mathbf{v})} \right]^2 \left[\frac{1}{20} - \frac{\tau_F(\mathbf{v})}{42\tau_0(\mathbf{v})} \right] & \tau_0(\mathbf{v}) \geq \tau_F(\mathbf{v}) \\ \frac{1}{12} - \frac{2}{35} \sqrt{\frac{\tau_0(\mathbf{v})}{\tau_F(\mathbf{v})}} & \tau_0(\mathbf{v}) < \tau_F(\mathbf{v}). \end{cases} \quad (4.37)$$

Let us choose T as per Eq. (4.29), with \mathcal{N} being the average flux. Now $\tau_0(\mathbf{v})$ is independent of $\mathcal{N}(t_0)$ (it only depends on \mathcal{N}), whereas $\tau_F(\mathbf{v})$ depends on $\mathcal{N}(t_0)$. The expression on the right-hand side of Eq. (4.37) is a convex function of

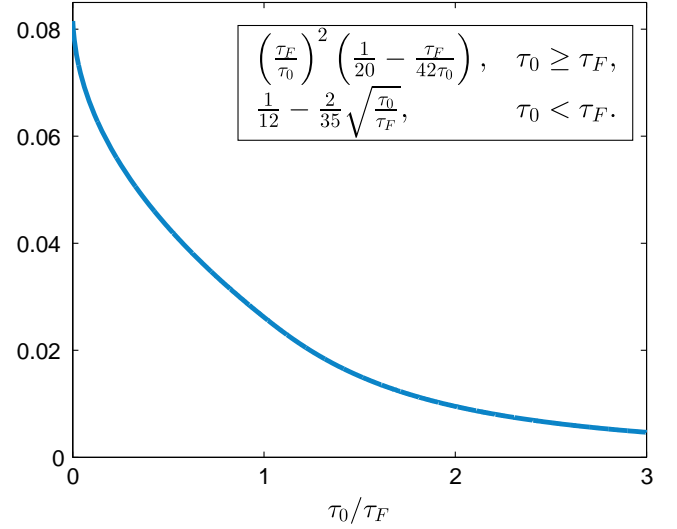


FIG. 5. A convex function that appears in Eq. (4.37).

$1/\tau_F(\mathbf{v})$, as plotted in Fig. 5. Therefore, using Jensen's inequality, we can place a lower bound on this expression by replacing $1/\tau_F(\mathbf{v})$ with its average value. As $1/\tau_F(\mathbf{v})$ is proportional to $\mathcal{N}(t_0)$, its average value can be found by just replacing $\mathcal{N}(t_0)$ with \mathcal{N} . As a result, we obtain a lower bound on $\bar{\Sigma}$ in terms of the time average of $\mathcal{N}(t_0)$, i.e., in terms of \mathcal{N} :

$$\bar{\Sigma} \geq c_Z \left(\frac{\kappa}{\mathcal{N}} \right)^{2(p-1)/(p+1)} \left[1 + c_\gamma \left(\frac{\kappa}{\mathcal{N}} \right)^{2p/(p+1)} \right]^{-1}. \quad (4.38)$$

In other words, we have a lower bound on the time-averaged mean-square error in terms of the time-averaged flux.

As before, for large \mathcal{N} the expression in square brackets is approximately 1, and we have

$$\bar{\Sigma} \geq c_Z \left(\frac{\kappa}{\mathcal{N}} \right)^{2(p-1)/(p+1)}. \quad (4.39)$$

Again, this result is not dependent on the particular means used in Eq. (4.21) to avoid a singularity at $\omega = 0$, as shown Appendix C.

V. ACHIEVING THE OPTIMAL SCALING

A lower bound is not a Heisenberg limit, and is of limited value, if it is not close to a realizable error. Here, we demonstrate that the scaling in Eq. (4.39) is indeed achievable in principle, up to a constant prefactor. The measurement technique we present is only intended to be possible in principle (allowed by the laws of quantum mechanics), analogous to the canonical measurements with optimal states for measurement of a constant phase [25]. It is presented because it enables us to prove rigorous results.

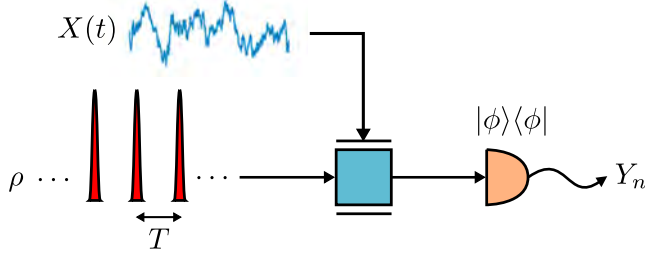


FIG. 6. A pulsed phase measurement scheme to achieve the optimal scaling.

How to perform measurements that would be experimentally realizable is a question for future work.

Consider an estimation strategy where the probe field is concentrated into pulses separated by time T , as shown in Fig. 6. This T is, as yet, unrelated to the T used to obtain the lower bound in the preceding section. Each pulse is assumed to be so short that the phase $X(t)$ does not vary significantly during the pulse duration. An average flux \mathcal{N} is ensured by each pulse having average photon number $\mathcal{N}T$.

We first assume that the phase modulation is weak; viz.,

$$\mathbb{E}[X^2(t)] \ll 1. \quad (5.1)$$

Using canonical phase measurements and minimum-uncertainty states within each pulse, the observation $Y_n \in (-\pi, \pi]$ at each sampling can be linearized as

$$Y_n \approx X(nT) + \xi_n, \quad (5.2)$$

where ξ_n is a random variable describing the noise with moments

$$\mathbb{E}(\xi_n|X) \approx 0, \quad \mathbb{E}(\xi_n \xi_m|X) \approx \delta_{nm} \frac{(4/27)|z_A|^3}{(\mathcal{N}T)^2}, \quad (5.3)$$

with $z_A \approx -2.338$ being the first negative root of the Airy function [64]. The above moments are exact in the asymptotic limit of large $\mathcal{N}T$.

The condition given by Eq. (5.1) can be relaxed for large phase fluctuations by making the canonical phase measurements adaptive [65], as shown in Appendix D. A rigorous accounting of the error due to phase ambiguity will be presented in Sec. VII in the case of a periodic distortion function. For the remainder of this section we assume Eqs. (5.1)–(5.3) for simplicity.

After all measurements are made, a final estimate of $X(t)$ can be constructed via the Whittaker-Shannon interpolation formula:

$$\begin{aligned} \check{X}(t) &:= \sum_{n=-\infty}^{\infty} Y_n \operatorname{sinc}\left(\frac{\pi t}{T} - \pi n\right) \\ &= X_T(t) + \xi(t), \end{aligned} \quad (5.4)$$

$$X_T(t) := \sum_{n=-\infty}^{\infty} X(nT) \operatorname{sinc}\left(\frac{\pi t}{T} - \pi n\right), \quad (5.5)$$

$$\xi(t) := \sum_{n=-\infty}^{\infty} \xi_n \operatorname{sinc}\left(\frac{\pi t}{T} - \pi n\right). \quad (5.6)$$

The mean-square error becomes

$$\begin{aligned} \mathbb{E}[[\check{X}(t) - X(t)]^2] &= \mathbb{E}[[\xi(t) + X_T(t) - X(t)]^2] \\ &= \mathbb{E}[\xi^2(t)] + \mathbb{E}[[X_T(t) - X(t)]^2], \end{aligned} \quad (5.7)$$

which consists of an aliasing error and a measurement error. Here, we have used the relation

$$\mathbb{E}[\xi(t)[X_T(t) - X(t)]] = 0. \quad (5.8)$$

As in the preceding section, we are interested in the limit of asymptotically large \mathcal{N} , which yields lower error for the estimation. Because the aliasing error will be smaller as the time between samples T is smaller, we consider the limit of small T . Averaging over time, and taking $\gamma T \ll \pi$, the aliasing error is

$$\begin{aligned} \frac{1}{T} \int_0^T dt \mathbb{E}[[X_T(t) - X(t)]^2] &= \frac{2}{\pi} \int_{\pi/T}^{\infty} d\omega \tilde{\Sigma}_0(\omega) \\ &\approx \frac{2(\kappa T)^{p-1}}{\pi^p(p-1)}, \end{aligned} \quad (5.9)$$

as shown in Appendix E. The measurement error is

$$\frac{1}{T} \int_0^T dt \mathbb{E}[\xi^2(t)] \approx \frac{(4/27)|z_A|^3}{(\mathcal{N}T)^2} \quad (5.10)$$

from Eq. (5.3). The overall error is hence

$$\bar{\Sigma} \approx \frac{2(\kappa T)^{p-1}}{\pi^p(p-1)} + \frac{(4/27)|z_A|^3}{(\mathcal{N}T)^2} \quad (5.11)$$

for $\gamma T \ll \pi$. Note that the first term increases with T , whereas the second term decreases with T . This result is as we expect, because increasing T means that the phase is sampled less frequently and can vary more in between samples, but increasing T also means that more power is available to estimate each sample, which reduces the error.

The optimal value of T , corresponding to minimizing the right-hand side of Eq. (5.11), is

$$T = \left(\frac{(4/27)|z_A|^3 \pi^p}{\mathcal{N}^2 \kappa^{p-1}} \right)^{1/(p+1)}, \quad (5.12)$$

which decreases with \mathcal{N} as we expect. It is slightly different from the T in the preceding section, but the scaling is the same. This calculation thus gives some insight into the T

used to define the function $v(t)$ in the preceding section: It relates to the correlation time for the intensity fluctuations in a light beam that will optimally probe the time-varying phase.

With this value of T , we obtain an average mean-square error

$$\bar{\Sigma} \approx c_A (\kappa/\mathcal{N})^{2(p-1)/(p+1)}, \quad (5.13)$$

where $c_A > c_Z$ is the dimensionless constant

$$c_A := \frac{p+1}{p-1} (4|z_A|^3/27)^{(p-1)/(p+1)} \pi^{-2p/(p+1)}. \quad (5.14)$$

Thus, the achievable variance has the same scaling with respect to κ/\mathcal{N} as that in the lower bound in Eq. (4.39) but with a larger prefactor. This demonstrates that the scaling of the lower bound in Eq. (4.39) with \mathcal{N} is tight, up to a constant prefactor, and thus represents a rigorous Heisenberg limit.

VI. PERIODIC QUANTUM ZIV-ZAKAI BOUND

A. Periodic distortion functions

We have considered phase estimation as an example of the application of the quantum Bell-Ziv-Zakai bounds. Phase measurements are intrinsically modulo 2π because they are unable to distinguish between phases that differ by multiples of 2π . For this reason, the phase will typically be taken to be in some standard region, such as $(-\pi, \pi]$. Then, a phase of $-\pi + \delta_1$, for some small $\delta_1 > 0$, can easily be estimated as $\pi - \delta_2$, for some $\delta_2 > 0$. It seems unrealistic to quantify the error as $\approx 2\pi$, because the phase difference is small modulo 2π . For this reason, it is better to use periodic distortion functions for measurements of this type.

A periodic distortion function satisfies

$$D(z + 2\pi m) = D(z), \quad (6.1)$$

for any real number z and integer m . We require the distortion function to satisfy conditions similar to those for the nonperiodic case: It should be symmetric, have $D(0) = 0$, and be differentiable and nondecreasing on $[0, \pi]$. A further requirement is that

$$\dot{D}(z) \geq \dot{D}(\pi - z), \quad (6.2)$$

for $\pi/2 \leq z \leq \pi$. This is a technical condition needed for the classical Ziv-Zakai bound of Ref. [54]. An example of a periodic distortion function satisfying these conditions is the periodic modification of the mean-square error,

$$D(z) = ([z]_{2\pi})^2, \quad (6.3)$$

where the notation

$$[z]_{2\pi} := z + 2\pi \text{floor}\left(\frac{1}{2} - \frac{z}{2\pi}\right) \quad (6.4)$$

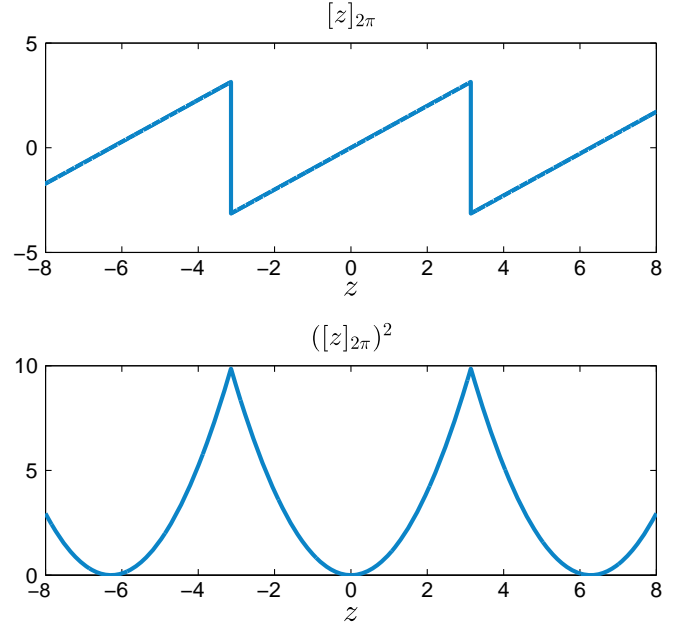


FIG. 7. Top panel: The modulo function $[z]_{2\pi}$. Bottom panel: Square of the modulo function.

denotes the value of z modulo the interval $(-\pi, \pi]$, as shown in Fig. 7.

Rather than considering the distortion function $D(\mathbf{u}^\top \boldsymbol{\epsilon})$ for a general vector \mathbf{u} , as in Sec. II, for simplicity, we restrict our attention to the case $u_\ell = \delta_{\ell,j}$. This corresponds to the distortion function $D(\epsilon_j)$, where ϵ_j denotes the j th component of the estimation error $\boldsymbol{\epsilon}$. This case is sufficient for bounding the average mean-square error of the estimate.

B. Periodic Ziv-Zakai bounds

In this subsection, we summarize the technique for obtaining periodic Ziv-Zakai bounds from Ref. [54]. One can obtain an expression similar to Eq. (2.8) as

$$\begin{aligned} \mathbb{E}[D(\epsilon_j)] &= \int_0^\pi dz D(z) P_{|[\epsilon_j]_{2\pi} |}(z) \\ &= [-D(z) \Pr(|[\epsilon_j]_{2\pi} | \geq z)] \Big|_0^\pi \\ &\quad + \int_0^\pi dz \dot{D}(z) \Pr(|[\epsilon_j]_{2\pi} | \geq z) \\ &= \frac{1}{2} \int_0^{2\pi} d\tau \dot{D}\left(\frac{\tau}{2}\right) \Pr\left(|[\epsilon_j]_{2\pi} | \geq \frac{\tau}{2}\right). \end{aligned} \quad (6.5)$$

This result is equivalent to Eq. (15) of Ref. [54]. Using the property (6.2) of D , Eq. (6.5) gives (see Appendix F)

$$\begin{aligned} \mathbb{E}[D(\epsilon_j)] &= \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \int_{-\pi}^\pi dx [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] \\ &\quad \times G(\pi_0, \mathbf{x}, \mathbf{v}, \tau), \end{aligned} \quad (6.6)$$

where \mathbf{v} is a vector as in Sec. II A, and

$$\begin{aligned}
 G(\pi_0, \mathbf{x}, \mathbf{v}, \tau) & \\
 := \pi_0 \Pr & \left(\left([\epsilon_j]_{2\pi} > \frac{\tau}{2} \right) \vee \left([\epsilon_j]_{2\pi} \leq \frac{\tau}{2} - \pi \right) \middle| \mathbf{X} = \mathbf{x} \right) \\
 & + \pi_1 \Pr \left(\left([\epsilon_j]_{2\pi} \leq -\frac{\tau}{2} \right) \right. \\
 & \left. \vee \left([\epsilon_j]_{2\pi} > \pi - \frac{\tau}{2} \right) \middle| \mathbf{X} = [\mathbf{x} + \mathbf{v}\tau]_{2\pi} \right). \quad (6.7)
 \end{aligned}$$

Similarly to Eqs. (2.12) and (2.13), the probabilities π_0 and π_1 are given by

$$\pi_0 := \frac{P_X(\mathbf{x})}{P_X(\mathbf{x}) + P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})}, \quad \pi_1 := 1 - \pi_0. \quad (6.8)$$

Now we require the vector \mathbf{v} to have its j component equal to 1, i.e., $v_j = 1$. This requirement corresponds to the condition $\mathbf{u}^\top \mathbf{v} = 1$ used in previous sections, for the special case $u_\ell = \delta_{\ell,j}$. With this restriction, consider a hypothesis-testing problem for distinguishing \mathbf{x} from $\mathbf{x} + \mathbf{v}\tau$. It is easily seen that the condition $([\epsilon_j]_{2\pi} > \tau/2) \vee ([\epsilon_j]_{2\pi} \leq \tau/2 - \pi)$ in the second line of Eq. (6.7) corresponds to the estimate $\check{X}_j(\mathbf{Y})$ being closer to $x_j + \tau$ than to x_j (see Fig. 3 of Ref. [54]). Similarly, the condition $([\epsilon_j]_{2\pi} \leq -\tau/2) \vee ([\epsilon_j]_{2\pi} > \pi - \tau/2)$ corresponds to $\check{X}_j(\mathbf{Y})$ being closer to x_j than to $x_j + \tau$ (see Fig. 4 of Ref. [54]). Hence, $G(\pi_0, \mathbf{x}, \mathbf{v}, \tau)$ is equal to the probability of error for distinguishing \mathbf{x} and $\mathbf{x} + \mathbf{v}\tau$ for the decision scheme based on whether x_j or $x_j + \tau$ is closer to the estimate $\check{X}_j(\mathbf{Y})$. It is therefore lower bounded by the error for an optimal decision scheme, which, together with Eq. (6.6), implies

$$\begin{aligned}
 \mathbb{E}[D(\epsilon_j)] & \geq \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \sup_{\mathbf{v}: v_j=1} \int_{-\pi}^\pi d\mathbf{x} [P_X(\mathbf{x}) \\
 & + P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})] P_e(\mathbf{x}, [\mathbf{x} + \mathbf{v}\tau]_{2\pi}). \quad (6.9)
 \end{aligned}$$

Using $\min\{P_X(\mathbf{x}), P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})\}$, similarly to the approach in Sec. II A, gives (see Appendix F)

$$\begin{aligned}
 \mathbb{E}[D(\epsilon_j)] & \geq \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \int_{-\pi}^\pi d\mathbf{x} \min[P_X(\mathbf{x}), \\
 & P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})] G(1/2, \mathbf{x}, \mathbf{v}, \tau). \quad (6.10)
 \end{aligned}$$

Here, $G(1/2, \mathbf{x}, \mathbf{v}, \tau)$ is the error probability for a decision problem with equal prior likelihood. Lower bounding the error probability then gives

$$\begin{aligned}
 \mathbb{E}[D(\epsilon_j)] & \geq \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \sup_{\mathbf{v}: v_j=1} \int_{-\pi}^\pi d\mathbf{x} \min[P_X(\mathbf{x}), \\
 & P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})] P_e^{\text{el}}(\mathbf{x}, [\mathbf{x} + \mathbf{v}\tau]_{2\pi}). \quad (6.11)
 \end{aligned}$$

This concludes the summary of the results in Ref. [54]. Next, using Eq. (2.23), for the case of quantum measurements, we finally obtain the periodic quantum Ziv-Zakai bound

$$\begin{aligned}
 \mathbb{E}[D(\epsilon_j)] & \geq \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \sup_{\mathbf{v}: v_j=1} \int_{-\pi}^\pi d\mathbf{x} \min[P_X(\mathbf{x}), \\
 & P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})] \left[1 - \sqrt{1 - F(\mathbf{x}, [\mathbf{x} + \mathbf{v}\tau]_{2\pi})} \right]. \quad (6.12)
 \end{aligned}$$

C. Gaussian prior

To provide a result for Gaussian variation of \mathbf{x} , we encounter a difficulty. Gaussian distributions always extend from $-\infty$ to $+\infty$ (even though they exponentially decay), whereas the variation of x_j is limited to the interval $(-\pi, \pi]$. The method we use is to take a Gaussian probability distribution P_G and wrap it around 2π . This would physically correspond to a case where a phase shift is caused by variation in an unbounded quantity (such as the position of a mirror on which the beam is incident [44]) that has Gaussian statistics. The corresponding periodic probability density is then given by

$$P_X(\mathbf{x}) = \sum_{\mathbf{n}} P_G(\mathbf{x} + 2\pi\mathbf{n}), \quad (6.13)$$

where the sum is over all vectors of integers, \mathbf{n} .

Denoting

$$H(\mathbf{x}) := \left[1 - \sqrt{1 - F(\mathbf{x}, [\mathbf{x} + \mathbf{v}\tau]_{2\pi})} \right], \quad (6.14)$$

the integral over \mathbf{x} in Eq. [23] can be lower bounded as

$$\begin{aligned}
 & \int_{-\pi}^\pi d\mathbf{x} \min[P_X(\mathbf{x}), P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})] H(\mathbf{x}) \\
 & = \int_{-\pi}^\pi d\mathbf{x} \min \left[\sum_{\mathbf{n}} P_G(\mathbf{x} + 2\pi\mathbf{n}), \right. \\
 & \quad \left. \sum_{\mathbf{m}} P_G(\mathbf{x} + 2\pi\mathbf{m} + \mathbf{v}\tau) \right] H(\mathbf{x}) \\
 & \geq \int_{-\pi}^\pi d\mathbf{x} \sum_{\mathbf{n}} \min[P_G(\mathbf{x} + 2\pi\mathbf{n}), H(\mathbf{x} + 2\pi\mathbf{n} + \mathbf{v}\tau)] H(\mathbf{x}) \\
 & = \int_{-\infty}^\infty d\mathbf{x} \min[P_G(\mathbf{x}), P_G(\mathbf{x} + \mathbf{v}\tau)] H([\mathbf{x}]_{2\pi}). \quad (6.15)
 \end{aligned}$$

We therefore obtain a result closely analogous to Eq. (2.29):

$$\begin{aligned}
 \mathbb{E}[D(\epsilon_j)] & \geq \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \\
 & \times \sup_{\mathbf{v}: v_j=1} \int d\mathbf{x} \min[P_G(\mathbf{x}), P_G(\mathbf{x} + \mathbf{v}\tau)] \\
 & \times \left[1 - \sqrt{1 - F([\mathbf{x}]_{2\pi}, [\mathbf{x} + \mathbf{v}\tau]_{2\pi})} \right], \quad (6.16)
 \end{aligned}$$

where the second integral is over *all* \mathbf{x} . On the right-hand side, the main difference between this expression and that in Eq. (2.29) is that the integral is up to $\tau = \pi$, rather than $\tau \rightarrow \infty$.

Next, we choose the mean-square error distortion function in Eq. (6.3) for periodic variables. Assuming generators that commute and satisfy Eq. (2.31) with addition modulo 2π , it follows via Eq. (6.16) above that the quantum Ziv-Zakai bound in Eq. (3.7) is modified to

$$\begin{aligned} \mathbb{E}[D(\epsilon_j)] &\geq \frac{1}{2} \sup_{\mathbf{v}: v_j=1} \int_0^\pi d\tau \tau \Lambda\left(\frac{\tau}{\tau_0(\mathbf{v})}\right) \Lambda\left(\sqrt{\frac{\tau}{\tau_F(\mathbf{v})}}\right) \\ &= \sup_{\mathbf{v}: v_j=1} Z_\pi(\mathbf{v}), \end{aligned} \quad (6.17)$$

where

$$Z_\pi(\mathbf{v}) := \begin{cases} \tau_F^2 \left(\frac{1}{20} - \frac{\tau_F(\mathbf{v})}{42\tau_0(\mathbf{v})} \right) & \pi \geq \tau_F(\mathbf{v}) \leq \tau_0(\mathbf{v}) \\ \tau_0^2 \left(\frac{1}{12} - \frac{2}{35} \sqrt{\frac{\tau_0(\mathbf{v})}{\tau_F(\mathbf{v})}} \right) & \tau_F(\mathbf{v}) > \tau_0(\mathbf{v}) \leq \pi \\ \pi^2 \left(\frac{1}{4} - \frac{\pi}{6\tau_0(\mathbf{v})} - \frac{\sqrt{\pi}}{5\sqrt{\tau_F(\mathbf{v})}} \right) & \tau_F(\mathbf{v}) > \pi < \tau_0(\mathbf{v}). \\ + \frac{\pi^{3/2}}{7\tau_0(\mathbf{v})\sqrt{\tau_F(\mathbf{v})}} & \end{cases} \quad (6.18)$$

A minor difference between this result and what we obtained before is that the left-hand side in Eq. (6.17) is the mean-square error (modulo 2π) for a single component for the error, whereas the left-hand side in Eq. (3.7) contains the full covariance matrix. In other words, Eq. (6.17) corresponds to taking $u_\ell = \delta_{\ell,j}$, to give the mean-square error for x_j . However, this is exactly what is used in Sec. IV. Therefore, we can follow an analysis similar to that in Sec. IV.

First, consider the case of constant flux; then, using Eq. (4.29) for T , we have $\tau_0(\mathbf{v}) < \tau_F(\mathbf{v})$. Next, we need either the second or third subfunction in Eq. (6.18). The third subfunction will be at least $\pi^2 11/420$ when the condition is satisfied. If $\tau_0 > \pi$, but the second subfunction is less than $\pi^2 11/420$, then the second subfunction still provides a lower bound on $Z_\pi(\mathbf{v})$. The net result is that we can lower bound $Z_\pi(\mathbf{v})$ by the minimum of $\pi^2 11/420$ and the second subfunction, and the lower bound on the variance is only slightly modified to

$$\begin{aligned} \Sigma(t_0, t_0) &\geq \min \left\{ \pi^2 \frac{11}{420}, \right. \\ &\quad \left. c_Z \left(\frac{\kappa}{\mathcal{N}} \right)^{2(p-1)/(p+1)} \left[1 + c_\gamma \left(\frac{\kappa}{\mathcal{N}} \right)^{2p/(p+1)} \right]^{-1} \right\}. \end{aligned} \quad (6.19)$$

In the case of time-dependent flux, we can again show that the lower bound in Eq. (6.18) is a convex function of

$1/\tau_F(\mathbf{v})$. This means that, for flux that is dependent on time, we again obtain a lower bound on the average value by replacing the instantaneous flux with the time-averaged flux \mathcal{N} . As a result, we obtain exactly the same lower bound on the time-averaged variance $\bar{\Sigma}$. In the limit of large \mathcal{N} , the lower bound on the variance is less than $\pi^2 11/420$. Thus, the min is not needed, and we obtain exactly the same asymptotic expression for the bound as we did in Sec. IV.

VII. ACHIEVING THE OPTIMAL SCALING: EFFECT OF PHASE AMBIGUITY

The analysis of the technique for achieving the optimal scaling given in Sec. V does not fully address the fact that the phase can only be measured modulo 2π . When tracking a phase, it is possible to resolve this ambiguity from the fact that the variation of the phase is continuous. Provided the phase does not change too much between successive estimates, and each estimate is reasonably accurate, we can keep track of changes by 2π . In other words, one can add suitable multiples of 2π to Y_n to give \check{X}_n , such that $|\check{X}_n - \check{X}_{n-1}| \leq \pi$. If the initial range of the phase is known, then the error should not exceed π .

This approach is problematic when the phase can vary arbitrarily far from zero, such as for a Wiener process. There is a nonzero possibility, however small, of choosing the wrong multiple of 2π at any step, and from that point on, there will continue to be an error of size 2π because of this initial error. This is called a phase-wrap error. When averaging measurements over an arbitrarily long period of time, the phase error can grow to be arbitrarily large. For the phase variation we consider, the Fourier spectrum $\tilde{\Sigma}_0(\omega)$ is bounded for $\omega = 0$, so the prior distribution has a bounded variance for any given time. This means that the error due to phase wraps is not unbounded, but it is still problematic.

Here, we consider the periodic distortion function in Eq. (6.3), given by the mean-square error modulo 2π . In this case, phase-wrap errors for individual points on their own do not matter because they do not increase a periodic distortion function. The problem appears when we consider estimation of the phase *between* the sample points, where the phase is interpolated. The estimated interpolation error for the Whittaker-Shannon interpolation formula is only accurate if there are no phase-wrap errors.

To simplify the problem, we take $X(t)$ to vary over the entire real line. Since the error is quantified modulo 2π , the estimation problem is identical to that where $X(t)$ is limited to the region $(-\pi, \pi]$. There is the additional advantage that the probability distribution is now exactly Gaussian, rather than given by Eq. (6.13).

In this section, let t be the particular time for which we wish to bound the mean-square deviation (modulo 2π) between the interpolated estimate and the actual phase. Let n_t be the sample time nearest t , in the sense that

$$n_t := \text{round}(t/T), \quad (7.1)$$

where T is the pulse separation. We use the ‘‘round half down’’ convention, so if t is equidistant between two sample times, we take the smaller sample time. Then, \check{X}_{n_t} is the phase estimate for the sample time nearest t .

Let L_n be the multiple of 2π that is added to the measurement result $Y_n \in (-\pi, \pi]$ to give \check{X}_n , so that

$$\check{X}_n = Y_n + 2\pi L_n. \quad (7.2)$$

We choose $L_{n_t} = 0$. Thus, \check{X}_{n_t} is just Y_{n_t} , and it is in the interval $(-\pi, \pi]$. In this section, the noise random variable ξ_n is redefined as

$$\xi_n := [Y_n - X(nT)]_{2\pi}. \quad (7.3)$$

With this definition, the moments given in Eq. (5.3) are correct. We then define an integer K_n as

$$K_n := \frac{1}{2\pi} [\check{X}_n - X(nT) - \xi_n], \quad (7.4)$$

so that

$$\check{X}_n = X(nT) + \xi_n + 2\pi K_n. \quad (7.5)$$

We also define $K := K_{n_t}$. The number of phase-wrap errors between samples is quantified by

$$Z_n := K_n - K_{n-1}. \quad (7.6)$$

The quantity K is the number of phase wrappings between \check{X}_{n_t} and $X(n_t T)$. This difference, in itself, is unimportant if deviations are only measured modulo 2π . What is important is that this difference is maintained for the other estimates. To achieve this result, we add or subtract multiples of 2π as needed (i.e., we choose L_n) to make the differences between neighboring estimates no more than π ; in particular, $\check{X}_n - \check{X}_{n-1} \in (-\pi, \pi]$. Provided certain conditions are met (discussed below), the difference between \check{X}_n and $X(nT)$ will be close to $2\pi K$ for all n , in other words, to the *same* multiple of 2π . Otherwise, there are phase-wrap errors, so K_n is dependent on n . This dependence on n will introduce error to the interpolation; however, this error can be bounded.

The interpolated values $\check{X}(t)$ can be expressed as

$$\check{X}(t) = X_T(t) + \xi(t) + 2\pi K(t), \quad (7.7)$$

where $X_T(t)$ and $\xi(t)$ are defined as in Eqs. (5.4) and (5.5), and

$$K(t) := \sum_{n=-\infty}^{\infty} K_n \text{sinc}\left(\frac{\pi t}{T} - \pi n\right). \quad (7.8)$$

This expression shows why we want the difference between the estimates and values to remain the same multiple of 2π . If the difference is the same, then $K(t)$ simply becomes the constant K . On the other hand, if K_n varies with n , then $K(t)$ will not be an integer, and the extra term $2\pi K(t)$ in Eq. (7.7) will give an increased error modulo 2π .

The analysis of the error in Eq. (5.7) now needs to be performed modulo 2π . First, define

$$\Delta(t) := [\xi(t) + X_T(t) - X(t)]_{2\pi}. \quad (7.9)$$

In terms of this quantity, we obtain

$$\begin{aligned} \mathbb{E}[[\check{X}(t) - X(t)]_{2\pi}^2] &= \mathbb{E}[[2\pi K(t) + \Delta(t)]_{2\pi}^2] \\ &\leq \mathbb{E}[\{[2\pi K(t)]_{2\pi} + \Delta(t)\}^2] \\ &= \mathbb{E}[[2\pi K(t)]_{2\pi}^2] + \mathbb{E}[\Delta^2(t)] + 2\mathbb{E}[\Delta(t)[2\pi K(t)]_{2\pi}] \\ &\leq \left\{ \sqrt{\mathbb{E}[[2\pi K(t)]_{2\pi}^2]} + \sqrt{\mathbb{E}[\Delta^2(t)]} \right\}^2. \end{aligned} \quad (7.10)$$

Note that we can upper bound the time-averaged value of $\mathbb{E}[[\check{X}(t) - X(t)]_{2\pi}^2]$ using the time-averaged values of $\mathbb{E}[\Delta^2(t)]$ and $\mathbb{E}[[2\pi K(t)]_{2\pi}^2]$, as per the last line above, because the expression on the last line is a concave function of these quantities. The time-averaged value of $\mathbb{E}[\Delta^2(t)]$ is exactly what was obtained in Sec. V, with the result given in Eq. (5.11). The remaining task is therefore to find the time-averaged value of $\mathbb{E}[[2\pi K(t)]_{2\pi}^2]$.

To achieve this, we need to bound the probabilities of phase-wrap errors. If t is a sample time so that $t = n_t T$, then $K(t) = K$, and $[2\pi K(t)]_{2\pi}$ is equal to zero. Therefore, in the remainder of this discussion, we assume that t is not a sample time. We wish to consider the error in interpolating at the given time t . Without loss of generality, this time can be taken to be in the interval $(0, T/2]$. This is because there is translation symmetry and time-reversal symmetry. One can simply translate the time by a multiple of T and change the sample numbering such that $n = 0$ or $n = 1$ corresponds to the closest sample time. That would yield $t \in (0, T)$. Then, if $t \in (T/2, T)$, one can simply reverse all times about $T/2$, so $t \in (0, T/2]$.

Since $n_t := \text{round}(t/T)$, for $t \in (0, T/2]$, we have $n_t = 0$. Hence, we are starting with \check{X}_0 in the interval $(-\pi, \pi]$, so $\check{X}_0 = Y_0$ and $L_0 = 0$. Then, all other values of L_n are selected such that $|\check{X}_n - \check{X}_{n-1}| \leq \pi$. The goal of this is to ensure that the values of K_n are equal (or at least close) to K_0 .

Recall that the quantity Z_n , defined in Eq. (7.6) as the difference between successive values of K_n , quantifies the number of phase-wrap errors between samples. The probability of a phase-wrap error between samples is therefore given by

$$p_{\text{err}} := \Pr(|Z_m| > 0). \quad (7.11)$$

We first place an upper bound on p_{err} , and then bound $\mathbb{E}[[2\pi K(t)]_{2\pi}^2]$ in terms of p_{err} .

Using Eqs. (7.2) and (7.4), we obtain

$$\begin{aligned} & 2\pi|K_n - K_{n-1}| \\ &= |\check{X}_n - \check{X}_{n-1} - \xi_n + \xi_{n-1} - X(nT) + X((n-1)T)| \\ &\leq |\check{X}_n - \check{X}_{n-1}| + |\xi_n - \xi_{n-1}| + |X(nT) - X((n-1)T)| \\ &\leq \pi + |\xi_n - \xi_{n-1}| + |X(nT) - X((n-1)T)|. \end{aligned} \quad (7.12)$$

In the last line, we have used the fact that the values of L_n have been chosen such that $|\check{X}_n - \check{X}_{n-1}| \leq \pi$. Now, if $|\xi_n - \xi_{n-1}| + |X(nT) - X((n-1)T)|$ is *less* than π , then $2\pi|K_n - K_{n-1}| < 2\pi$. Because K_n takes integer values, this inequality implies that $K_n = K_{n-1}$, i.e., that there is no phase-wrap error between the two successive samples. In particular, there will be *no* phase-wrap error if $|\xi_n - \xi_{n-1}| < \pi/2$ and $|X(nT) - X((n-1)T)| < \pi/2$. This means that

$$\begin{aligned} p_{\text{err}} &\leq \Pr(|X(nT) - X((n-1)T)| \geq \pi/2) \\ &\quad + \Pr(|\xi_n - \xi_{n-1}| \geq \pi/2). \end{aligned} \quad (7.13)$$

To bound the first probability on the right-hand side in Eq. (7.13), we use the fact that the statistics for X are Gaussian. It turns out that $\mathbb{E}[[X(t) - X(t')]^2]$ can be bounded as a polynomial in $|t - t'|$ (see Appendix G). Here, $|t - t'| = T$, and we choose T to vary polynomially in κ/\mathcal{N} . As a result, $\mathbb{E}[[X(nT) - X((n-1)T)]^2]$ varies polynomially in κ/\mathcal{N} , becoming small for large \mathcal{N}/κ . Because the probability distribution for X is Gaussian, the probability of the difference between $X(nT)$ and $X((n-1)T)$ being larger than $\pi/2$ is exponentially small in \mathcal{N}/κ .

Next, we consider the probability of $|\xi_n - \xi_{n-1}|$ exceeding $\pi/2$. By modifying the phase measurements at each sample, it is possible to ensure that the probability of error greater than $\pi/4$ is exponentially small in \mathcal{N}/κ while not modifying the mean-square error to leading order (see Appendix H). Therefore, the probability of $|\xi_n - \xi_{n-1}|$ exceeding $\pi/2$ is exponentially small in \mathcal{N}/κ . Combining this with Eq. (7.13), and the fact that $\Pr(|X(nT) - X((n-1)T)| \geq \pi/2)$ is exponentially small in \mathcal{N}/κ , we find that p_{err} is exponentially small in \mathcal{N}/κ .

To upper bound $\mathbb{E}[[2\pi K(t)]_{2\pi}^2]$, we express K_n in terms of Z_n . We obtain, for positive n ,

$$K_n = K_0 + \sum_{m=1}^n Z_m. \quad (7.14)$$

Similarly, for negative n ,

$$K_n = K_0 - \sum_{m=n+1}^0 Z_m. \quad (7.15)$$

Therefore, we can write $K(t)$ as

$$\begin{aligned} K(t) &= K_0 + \sum_{n=1}^{\infty} \sum_{m=1}^n Z_m \text{sinc}\left(\frac{\pi t}{T} - \pi n\right) \\ &\quad + \sum_{n=-\infty}^{-1} \sum_{m=n+1}^0 Z_m \text{sinc}\left(\frac{\pi t}{T} - \pi n\right) \\ &= K_0 + \sum_{m=1}^{\infty} Z_m \sum_{n=m}^{\infty} \text{sinc}\left(\frac{\pi t}{T} - \pi n\right) \\ &\quad + \sum_{m=-\infty}^0 Z_m \sum_{n=-\infty}^{m-1} \text{sinc}\left(\frac{\pi t}{T} - \pi n\right) \\ &= K_0 + \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} Z_m a_m, \end{aligned} \quad (7.16)$$

where (see Appendix D)

$$\begin{aligned} a_m &:= (-1)^m \sin\left(\frac{\pi t}{T}\right) \\ &\quad \times \begin{cases} \Psi\left(\frac{m}{2} - \frac{t}{2T}\right) - \Psi\left(\frac{1}{2} + \frac{m}{2} - \frac{t}{2T}\right) & m > 0 \\ \Psi\left(\frac{1}{2} - \frac{m}{2} + \frac{t}{2T}\right) - \Psi\left(1 - \frac{m}{2} + \frac{t}{2T}\right) & m \leq 0, \end{cases} \end{aligned} \quad (7.17)$$

and Ψ is the digamma function $\Psi(x) = \Gamma'(x)/\Gamma(x)$.

Using this expression, we can bound $\mathbb{E}[[2\pi K(t)]_{2\pi}^2]$ as follows:

$$\begin{aligned} & \mathbb{E}[[2\pi K(t)]_{2\pi}^2] \\ &= \mathbb{E}[[2\pi(K(t) - K_0)]_{2\pi}^2] \\ &\leq \mathbb{E}[[2\pi(K(t) - K_0)]^2] \\ &= \mathbb{E}\left[\left(\sum_{n=-\infty}^{\infty} Z_n a_n\right)^2\right] \\ &= \sum_{m=-\infty}^{\infty} \mathbb{E}[Z_m^2] a_m^2 + 2 \sum_{m=-\infty}^{\infty} \mathbb{E}[Z_m Z_{m+1}] a_m a_{m+1} + S \\ &\lesssim p_{\text{err}} \left[\sum_{m=-\infty}^{\infty} a_m^2 + 2 \sum_{m=-\infty}^{\infty} |a_m a_{m+1}| \right] + S, \end{aligned} \quad (7.18)$$

where

$$S := \sum_{n,m: |n-m|>1} \mathbb{E}[Z_n Z_m] a_n a_m. \quad (7.19)$$

We use the notation \lesssim to indicate that higher-order terms in \mathcal{N} have been omitted. In the last line of Eq. (7.18), we have used the fact that Z_m is limited to $\{0, -1, +1\}$ with high probability, which implies $\mathbb{E}[Z_m^2] \approx p_{\text{err}}$ (see Appendix J).

In addition, $|\mathbb{E}[Z_m Z_{m+1}]|$ cannot exceed $\mathbb{E}[Z_m^2]$, which gives the inequality.

The quantity in square brackets in the last line of Eq. (7.18) is less than 1 (this can be seen numerically). Moreover, it can be shown that the longer-range correlations quantified by S are exponentially small (see Appendix J). Hence, we find that $\mathbb{E}[[2\pi K(t)]_{2\pi}^2]$ is exponentially small in \mathcal{N}/κ . This result is for the worst-case value of t (i.e., midway between two sample points), so averaging over t can only give smaller values.

Rather than deriving an optimal value of T , we simply use the value in Eq. (5.12). The scaling for the average variance given in Eq. (5.13) corresponds to the time average of $\mathbb{E}[\Delta^2(t)]$, with T given as in Eq. (5.12). Denoting this quantity by $\bar{\Sigma}_0$ and the time-averaged value of $\mathbb{E}[[2\pi K(t)]_{2\pi}^2]$ by $\bar{\Sigma}_{\text{wrap}}$, using Eq. (7.10) then gives

$$\bar{\Sigma} \lesssim \left(\sqrt{\bar{\Sigma}_{\text{wrap}}} + \sqrt{\bar{\Sigma}_0} \right)^2 \lesssim c_A (\kappa/\mathcal{N})^{2(p-1)/(p+1)}, \quad (7.20)$$

where c_A is defined in Eq. (5.14). Because $\bar{\Sigma}_{\text{wrap}}$ is of higher order in \mathcal{N} , it does not affect the final scaling constant for the leading-order term. Thus, we find that, when we fully take account of phase-wrap errors, we still obtain

$$\bar{\Sigma} = O((\kappa/\mathcal{N})^{2(p-1)/(p+1)}). \quad (7.21)$$

VIII. CONCLUSIONS

While fundamental quantum limits to the accuracy of estimation of single quantities are well known, deriving fundamental limits for estimating multiple parameters becomes very challenging when there is prior information describing correlations between them. A particularly important example of this is in phase estimation, where the phase at any time is correlated with the phase at earlier and later times. Such phase estimation tasks are needed, for example, in gravitational-wave astronomy.

Here, we have proven quantum forms of the Bell-Ziv-Zakai bounds for multiparameter estimation. One of these enables us to bound the accuracy possible when measuring a phase with stationary Gaussian prior statistics and a power-law spectrum $\sim \omega^{-p}$. We have thereby been able to prove that the scaling bound found in Ref. [41], for quantum states having time-symmetric stationary Gaussian statistics for the field quadratures, in fact holds for all possible quantum states.

Moreover, we have shown here analytically that the lower bound we have derived is always achievable, up to a constant prefactor. Specifically, it is possible, in principle, to achieve it by sampling with a regularly timed sequence of pulses, each of which is measured by a canonical phase measurement, and with interpolation of the phase between those times. This bound can therefore be regarded as analogous to the Heisenberg limit for measurement of a single constant phase.

We have also provided bounds for periodic distortion functions. An example of this is measurement of phase modulo 2π , so the mean-square error is evaluated modulo 2π . We find that the bounds we derive for the nonperiodic case hold almost unchanged.

In addition to providing the asymptotic form of the bounds for large flux, we have provided strict bounds that hold for all values of the flux. Our results provide a definite theoretical prediction, that the average mean-square error of any measurement technique cannot be below the bounds given. Another interesting prediction is that, for phase variation close to $1/f$ noise (i.e., $p \approx 1$), coherent states provide close to the best possible probe states.

The question of how to achieve these bounds is more difficult. The measurement technique we have described is possible in principle, in that it is allowed by the laws of quantum mechanics. A more practical scheme for measuring a varying phase would be to use a squeezed state and adaptive phase measurements, as has been described, for example, in Refs. [55,66], and experimentally demonstrated in Ref. [43]. Numerical results indicate that such techniques can give mean-square errors comparable to the lower bound for $p = 2$, but it is an open question how effective these techniques would be for other values of p .

Of course, all real measurements are limited by loss. We expect that real measurements would only give the scaling predicted up to some maximum flux dependent on the amount of loss and thereafter scale in the same way as for coherent states (given in Ref. [41]). The bound for measurements with loss is a question for future work. Some preliminary results on that question are given in Ref. [9].

Another important open question is how to reduce the gap between the upper and lower bounds on the minimum achievable variance. The ratio c_A/c_Z (of the prefactors for the upper and lower bounds) is about 14 to 20 for moderate values of p , but it is unbounded in the limit $p \rightarrow 1$ or $p \rightarrow \infty$. For the case of measurements of a constant phase, the exact minimum variance is known [25], and it would be desirable to achieve the same result here. Some possible approaches to this are to use a better function $v(t)$ (as we have not attempted to optimize this function), as well as to use different forms of the Ziv-Zakai bound.

We also note that while Gaussian correlations were assumed for the applied phase shift X [e.g., in Eq. (2.33)], we believe our method can be generalized to yield estimation bounds for a wide class of non-Gaussian correlations. More generally, there are many other multiparameter estimation tasks, in which there are prior constraints on the correlations, for which our quantum Bell-Ziv-Zakai bounds could reveal the ultimate achievable limits.

ACKNOWLEDGMENTS

Discussions with Ranjith Nair are gratefully acknowledged. D.W.B. is supported by ARC Grant No. FT100100761. M.T. is supported by the Singapore

National Research Foundation under NRF Grant No. NRF-NRFF2011-07. M. J. W. H. and H. M. W. are supported by the ARC Centre of Excellence Grant No. CE110001027.

APPENDIX A: VERSION OF BOUNDS WITH VALLEY-FILLING FUNCTION

Here, we provide slightly tighter forms of the quantum Bell-Ziv-Zakai bounds using the valley-filling function. The valley-filling function \mathcal{V} is defined by

$$\mathcal{V}\{f(\tau)\} := \sup_{\eta: \eta \geq 0} f(\tau + \eta). \quad (\text{A1})$$

Using Eqs. (31) and (35) of Ref. [52] to bound $\Pr(|\mathbf{u}^\top \boldsymbol{\epsilon}| \geq \tau/2)$ and noting Property 1 of Ref. [52] yields the Bell-Ziv-Zakai bounds [52,56]:

$$\begin{aligned} \mathbb{E}[D(\mathbf{u}^\top \boldsymbol{\epsilon})] &\geq \frac{1}{2} \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \mathcal{V}\left\{ \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int d\mathbf{x} \right. \\ &\quad \left. \times [P_X(\mathbf{x}) + P_X(\mathbf{x} + \mathbf{v}\tau)] P_e(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau) \right\}, \end{aligned} \quad (\text{A2})$$

$$\begin{aligned} &\geq \int_0^\infty d\tau \dot{D}\left(\frac{\tau}{2}\right) \mathcal{V}\left\{ \sup_{\mathbf{v}: \mathbf{u}^\top \mathbf{v}=1} \int d\mathbf{x} \right. \\ &\quad \left. \times \min[P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)] P_e^{\text{el}}(\mathbf{x}, \mathbf{x} + \mathbf{v}\tau) \right\}. \end{aligned} \quad (\text{A3})$$

These inequalities are tighter versions of Eqs. (2.17) and (2.20). The quantum bounds are obtained by lower bounding P_e and P_e^{el} using Eq. (2.22) or (2.23) with the right prior probabilities.

APPENDIX B: EVALUATION OF THE INTEGRAL

Here, we show how to evaluate the integral

$$\mathcal{I} := \int d\mathbf{x} \min[P_X(\mathbf{x}), P_X(\mathbf{x} + \mathbf{v}\tau)]. \quad (\text{B1})$$

Writing this integral out explicitly gives

$$\begin{aligned} \mathcal{I} &= \frac{1}{\sqrt{(2\pi)^k |\Sigma_0|}} \\ &\quad \times \int d\mathbf{x} \min \left[\exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma_0^{-1}(\mathbf{x} - \boldsymbol{\mu})\right), \right. \\ &\quad \left. \exp\left(-\frac{1}{2}(\mathbf{x} + \mathbf{v}\tau - \boldsymbol{\mu})^\top \Sigma_0^{-1}(\mathbf{x} + \mathbf{v}\tau - \boldsymbol{\mu})\right) \right], \end{aligned} \quad (\text{B2})$$

where k is the dimension of the vector and $\boldsymbol{\mu}$ is the mean of the distribution. Let V be an orthogonal matrix that diagonalizes Σ_0 , and D the matrix of eigenvalues of Σ_0 , so $\Sigma_0^{-1} = V^\top D^{-1} V$. Next, let U be an orthogonal matrix such that $\mathbf{w} := UD^{-1/2} V \mathbf{v}$ is a vector with only one nonzero entry. The magnitude of this vector is $|\mathbf{w}| = (\mathbf{v}^\top \Sigma_0^{-1} \mathbf{v})^{1/2}$. Taking a change of variables $\mathbf{z} = UD^{-1/2} V(\mathbf{x} - \boldsymbol{\mu})$, the integral becomes

$$\begin{aligned} \mathcal{I} &= \frac{1}{\sqrt{(2\pi)^k}} \int d\mathbf{z} \min \left[\exp\left(-\frac{1}{2} \mathbf{z}^\top \mathbf{z}\right), \right. \\ &\quad \left. \exp\left(-\frac{1}{2}(\mathbf{z} + \mathbf{w}\tau)^\top (\mathbf{z} + \mathbf{w}\tau)\right) \right]. \end{aligned} \quad (\text{B3})$$

Integrating each of the components of \mathbf{z} that are orthogonal to \mathbf{w} gives

$$\begin{aligned} \mathcal{I} &= \frac{1}{\sqrt{2\pi}} \int dz \min \left[\exp\left(-\frac{1}{2} z^2\right), \right. \\ &\quad \left. \exp\left(-\frac{1}{2}(z + |\mathbf{w}|\tau)^2\right) \right]. \end{aligned} \quad (\text{B4})$$

Splitting the integral up into the parts where each term in the minimum is smaller, we can then evaluate it to give

$$\begin{aligned} \mathcal{I} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-|\mathbf{w}|\tau/2} dz \exp\left(-\frac{1}{2} z^2\right) \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_{-|\mathbf{w}|\tau/2}^{\infty} dz \exp\left(-\frac{1}{2}(z + |\mathbf{w}|\tau)^2\right) \\ &= \frac{2}{\sqrt{2\pi}} \int_{|\mathbf{w}|\tau/2}^{\infty} dz \exp\left(-\frac{1}{2} z^2\right) \\ &= \text{erfc}\left(\frac{\sqrt{\pi}\tau}{2\tau_0(\mathbf{v})}\right), \end{aligned} \quad (\text{B5})$$

which yields the result in Eq. (2.33).

APPENDIX C: ASYMPTOTIC SCALING FOR SPECTRA WITH POWER-LAW TAIL

More generally, the prior power spectral density $\tilde{\Sigma}_0(\omega)$ may scale as $\kappa^{p-1}/|\omega|^p$ for $|\omega| \rightarrow \infty$. Making this concept rigorous, we assume that there exist constants w_0 and G such that

$$\tilde{\Sigma}_0(\omega) \geq \begin{cases} G/\kappa & |\omega| < \kappa w_0 \\ \kappa^{p-1}/|\omega|^p & |\omega| \geq \kappa w_0. \end{cases} \quad (\text{C1})$$

Then, we obtain, with $v(t)$ given in Eq. (4.23),

$$\begin{aligned}
 \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{|\tilde{v}(\omega)|^2}{\tilde{\Sigma}_0(\omega)} &= 2\pi T^2 \int_{-1/T}^{1/T} d\omega \frac{(1 - |\omega|T)^2}{\tilde{\Sigma}_0(\omega)} \\
 &\leq 4\pi T^2 \kappa \int_0^{\kappa w_0} d\omega \frac{(1 - \omega T)^2}{G} \\
 &\quad + 4\pi T^2 \int_0^{1/T} d\omega \frac{\omega^p (1 - \omega T)^2}{\kappa^{p-1}} \\
 &\leq \frac{4\pi T^2 \kappa^2 w_0}{G} + \frac{8\pi}{p_3 \kappa^{p-1} T^{p-1}}. \quad (\text{C2})
 \end{aligned}$$

The first term is negligible, provided that κT is small. With the expression we take for T , $\kappa T \propto [\kappa/\mathcal{N}(t_0)]^{2/(p+1)}$. Because we consider scaling with large $\mathcal{N}(t_0)/\kappa$, the first term is negligible, and we again obtain the result in the main text.

APPENDIX D: CANONICAL PHASE-LOCKED LOOP

The accuracy of our linear model in Sec. V relies on the assumption of weak phase modulation. For large phase fluctuations, we can borrow from the phase-locked loop concept [42–44,65–70] and modulate each pulse by an adaptive phase $-\tilde{X}(nT)$ before the canonical phase measurement, where $\tilde{X}(nT)$ is a causal estimate of $X(nT)$ extrapolated from previous observations $\{Y_{n-1}, Y_{n-2}, \dots\}$. Provided that $\tilde{X}(nT)$ tracks $X(nT)$ closely, viz.,

$$\mathbb{E}[[X(nT) - \tilde{X}(nT)]^2] \ll 1, \quad (\text{D1})$$

the net phase modulation $X(nT) - \tilde{X}(nT)$ will be small, and $Y_n \in (-\pi, \pi]$ can be linearized as

$$Y_n \approx X(nT) - \tilde{X}(nT) + \xi_n. \quad (\text{D2})$$

The requirement of small causal error according to Eq. (D1) is now less stringent than Eq. (5.1). To evaluate the causal error analytically, we approximate the discrete observations Y_n as a continuous-time signal given by

$$Y(t) \approx X(t) - \tilde{X}(t) + \xi(t), \quad (\text{D3})$$

$$\mathbb{E}[\xi(t)\xi(t')|X] = R\delta(t - t'), \quad (\text{D4})$$

$$R := \frac{(4/27)|z_A|^3}{\mathcal{N}^2 T}. \quad (\text{D5})$$

The continuous approximation is accurate in the high- \mathcal{N} limit because the measurement period T in Eq. (5.12) can be made arbitrarily small in the limit. The minimum causal error at a steady state is then given by the Yovits-Jackson formula [50]:

$$\mathbb{E}[[X(t) - \tilde{X}(t)]^2] \approx R \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \ln \left[1 + \frac{\tilde{\Sigma}_0(\omega)}{R} \right]. \quad (\text{D6})$$

Since the error decreases with decreasing R and increasing \mathcal{N} , the phase tracking can be made arbitrarily accurate in the high- \mathcal{N} limit. These considerations are similar to the principles of a homodyne phase-locked loop [66–69], except that here we assume canonical phase measurements to avoid photon-number fluctuations.

In the long-time limit, phase-wrap errors, no matter how rare, can still occur, making the estimate diverge from the true waveform by multiples of 2π . Just like the classical phase modulation system, it can be expected that this divergence will be eliminated by adding a dc notch filter to the output [69].

APPENDIX E: TIME AVERAGES

Here, we show how to take the time averages (5.9) and (5.10) given in the main text. Each average is taken because the error will depend on how far t is from the nearest sampling point. Because the distribution is otherwise time invariant, we need only average over the interval $[0, T]$.

Here, we take $S(t) := \Sigma_0(t, 0)$ and $f := 1/T$. Note that, because of stationary statistics, $\Sigma_0(t_1, t_2)$ depends only on $t_1 - t_2$. Evaluating Eq. (5.9) gives

$$\begin{aligned}
 \frac{1}{T} \int_0^T dt \mathbb{E}[X_T(t) - X(t)]^2 &= \frac{1}{T} \int_0^T dt \left[\sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \mathbb{E}[X(nT)X(mT)] \text{sinc}(\pi(ft - n)) \text{sinc}(\pi(ft - m)) \right. \\
 &\quad \left. + \mathbb{E}[X(t)X(t)] - 2 \sum_{n=-\infty}^{\infty} \mathbb{E}[X(nT)X(t)] \text{sinc}(\pi(ft - n)) \right] \\
 &= \frac{1}{T} \int_0^T dt \left[\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} S(nT) \text{sinc}(\pi(ft - n - m)) \text{sinc}(\pi(ft - m)) + S(0) \right. \\
 &\quad \left. - 2 \sum_{n=-\infty}^{\infty} S(t - nT) \text{sinc}(\pi(ft - n)) \right]
 \end{aligned}$$

$$\begin{aligned}
&= S(0) + \frac{1}{T} \int_{-\infty}^{\infty} dt \left[\sum_{n=-\infty}^{\infty} S(nT) \text{sinc}(\pi(ft-n)) \text{sinc}(\pi ft) - 2S(t) \text{sinc}(\pi ft) \right] \\
&= S(0) + \sum_{n=-\infty}^{\infty} S(nT) \text{sinc}(\pi n) - \frac{2}{T} \int_{-\infty}^{\infty} dt S(t) \text{sinc}(\pi ft) \\
&= 2S(0) - \frac{2}{T} \int_{-\infty}^{\infty} dt S(t) \text{sinc}(\pi ft) = 2S(0) - \frac{1}{\pi} \int_{-\pi f}^{\pi f} d\omega \int_{-\infty}^{\infty} dt S(t) e^{-i\omega t} \\
&= \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \tilde{\Sigma}_0(\omega) - \frac{1}{\pi} \int_{-\pi f}^{\pi f} d\omega \tilde{\Sigma}_0(\omega) = \frac{1}{\pi} \int_{\pi f}^{\infty} d\omega \tilde{\Sigma}_0(\omega) + \frac{1}{\pi} \int_{-\infty}^{-\pi f} d\omega \tilde{\Sigma}_0(\omega). \tag{E1}
\end{aligned}$$

Evaluating the integral for $\tilde{\Sigma}_0(\omega) = \kappa^{p-1}/(|\omega|^p + \gamma^p)$ gives

$$\frac{2}{\pi} \int_{\pi f}^{\infty} d\omega \tilde{\Sigma}_0(\omega) = \frac{2}{\pi} \int_{\pi f}^{\infty} d\omega \frac{\kappa^{p-1}}{\omega^p + \gamma^p} \approx \frac{2(\kappa T)^{p-1}}{\pi^p(p-1)}, \tag{E2}$$

assuming $\pi f = \pi/T \gg \gamma$.

Next, evaluating the average in Eq. (5.10) gives

$$\begin{aligned}
\frac{1}{T} \int_0^T dt \mathbb{E}[\xi^2(t)] &= \frac{1}{T} \int_0^T dt \sum_{n=-\infty}^{\infty} \mathbb{E}[\xi_n^2] \text{sinc}^2(\pi(ft-n)) \\
&\approx \frac{1}{T} \int_0^T dt \sum_{n=-\infty}^{\infty} \frac{(4/27)|z_A|^2}{(\mathcal{N}T)^2} \text{sinc}^2(\pi(ft-n)) \\
&= \frac{(4/27)|z_A|^2}{(\mathcal{N}T)^2} \frac{1}{T} \int_{-\infty}^{\infty} dt \text{sinc}^2(\pi ft) \\
&= \frac{(4/27)|z_A|^2}{(\mathcal{N}T)^2}. \tag{E3}
\end{aligned}$$

APPENDIX F: DERIVING THE BOUND IN THE PERIODIC CASE

Here, we show how to obtain Eq. (2.38) from Eq. (6.5), in order to provide a bound in the periodic case. Using the property (6.2) of D , Eq. (6.5) gives

$$\begin{aligned}
\mathbb{E}[D(\epsilon_j)] &= \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \Pr\left(|[\epsilon_j]_{2\pi}| \geq \frac{\tau}{2}\right) + \frac{1}{2} \int_\pi^{2\pi} d\tau \dot{D}\left(\frac{\tau}{2}\right) \Pr\left(|[\epsilon_j]_{2\pi}| \geq \frac{\tau}{2}\right) \\
&\geq \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \Pr\left(|[\epsilon_j]_{2\pi}| \geq \frac{\tau}{2}\right) + \frac{1}{2} \int_\pi^{2\pi} d\tau \dot{D}\left(\pi - \frac{\tau}{2}\right) \Pr\left(|[\epsilon_j]_{2\pi}| \geq \frac{\tau}{2}\right) \\
&= \frac{1}{2} \int_0^\pi d\tau \dot{D}\left(\frac{\tau}{2}\right) \left[\Pr\left(|[\epsilon_j]_{2\pi}| \geq \frac{\tau}{2}\right) + \Pr\left(|[\epsilon_j]_{2\pi}| \geq \pi - \frac{\tau}{2}\right) \right]. \tag{F1}
\end{aligned}$$

Introducing a vector \mathbf{v} as in Sec. II A, one can place a lower bound on the sum of probabilities in the expression above, via

$$\begin{aligned}
 & \Pr\left(|[\epsilon_j]_{2\pi}| \geq \frac{\tau}{2}\right) + \Pr\left(|[\epsilon_j]_{2\pi}| \geq \pi - \frac{\tau}{2}\right) \\
 &= \Pr\left([\epsilon_j]_{2\pi} > \frac{\tau}{2}\right) + \Pr\left([\epsilon_j]_{2\pi} \leq -\frac{\tau}{2}\right) + \Pr\left([\epsilon_j]_{2\pi} > \pi - \frac{\tau}{2}\right) + \Pr\left([\epsilon_j]_{2\pi} \leq \frac{\tau}{2} - \pi\right) \\
 &= \int_{-\pi}^{\pi} dx P_X(\mathbf{x}) \Pr\left([\epsilon_j]_{2\pi} > \frac{\tau}{2} \middle| \mathbf{X} = \mathbf{x}\right) + \int_{-\pi}^{\pi} dx P_X(\mathbf{x}) \Pr\left([\epsilon_j]_{2\pi} \leq -\frac{\tau}{2} \middle| \mathbf{X} = \mathbf{x}\right) \\
 &\quad + \int_{-\pi}^{\pi} dx P_X(\mathbf{x}) \Pr\left([\epsilon_j]_{2\pi} > \pi - \frac{\tau}{2} \middle| \mathbf{X} = \mathbf{x}\right) + \int_{-\pi}^{\pi} dx P_X(\mathbf{x}) \Pr\left([\epsilon_j]_{2\pi} \leq \frac{\tau}{2} - \pi \middle| \mathbf{X} = \mathbf{x}\right) \\
 &= \int_{-\pi}^{\pi} dx P_X(\mathbf{x}) \Pr\left([\epsilon_j]_{2\pi} > \frac{\tau}{2} \middle| \mathbf{X} = \mathbf{x}\right) + \int_{-\pi}^{\pi} dx P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi}) \Pr\left([\epsilon_j]_{2\pi} \leq -\frac{\tau}{2} \middle| \mathbf{X} = [\mathbf{x} + \mathbf{v}\tau]_{2\pi}\right) \\
 &\quad + \int_{-\pi}^{\pi} dx P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi}) \Pr\left([\epsilon_j]_{2\pi} > \pi - \frac{\tau}{2} \middle| \mathbf{X} = [\mathbf{x} + \mathbf{v}\tau]_{2\pi}\right) + \int_{-\pi}^{\pi} dx P_X(\mathbf{x}) \Pr\left([\epsilon_j]_{2\pi} \leq \frac{\tau}{2} - \pi \middle| \mathbf{X} = \mathbf{x}\right) \quad (\text{F2}) \\
 &= \int_{-\pi}^{\pi} dx [P_X(\mathbf{x}) + P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})] \left\{ \pi_0 \Pr\left(\left([\epsilon_j]_{2\pi} > \frac{\tau}{2}\right) \vee \left([\epsilon_j]_{2\pi} \leq \frac{\tau}{2} - \pi\right) \middle| \mathbf{X} = \mathbf{x}\right) \right. \\
 &\quad \left. + \pi_1 \Pr\left(\left([\epsilon_j]_{2\pi} \leq -\frac{\tau}{2}\right) \vee \left([\epsilon_j]_{2\pi} > \pi - \frac{\tau}{2}\right) \middle| \mathbf{X} = [\mathbf{x} + \mathbf{v}\tau]_{2\pi}\right) \right\}. \quad (\text{F3})
 \end{aligned}$$

The probabilities π_0 and π_1 are given by Eq. (6.8). Using this result in Eq. (F1) then implies Eq. (6.6).

Alternatively, one can replace $P_X(\mathbf{x})$ and $P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})$ with $\min\{P_X(\mathbf{x}), P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})\}$ in Eq. (F2), which yields

$$\begin{aligned}
 & \Pr\left(|[\epsilon_j]_{2\pi}| \geq \frac{\tau}{2}\right) + \Pr\left(|[\epsilon_j]_{2\pi}| \geq \pi - \frac{\tau}{2}\right) \\
 & \geq 2 \int_{-\pi}^{\pi} dx \min[P_X(\mathbf{x}), P_X([\mathbf{x} + \mathbf{v}\tau]_{2\pi})] G(1/2, \mathbf{x}, \mathbf{v}, \tau). \quad (\text{F4})
 \end{aligned}$$

Using this result in Eq. (F1) then implies Eq. (6.10).

APPENDIX G: BOUNDING CHANGE IN X

Here, we show how to bound $\mathbb{E}[|X(t) - X(t')|^2]$. In general, using only the property that $\tilde{\Sigma}_0(\omega)$ is an even function,

$$\begin{aligned}
 \mathbb{E}[|X(t) - X(t')|^2] &= 2[\Sigma_0(t, t) - \Sigma_0(t, t')] \\
 &= \int_{-\infty}^{\infty} \frac{d\omega}{\pi} \tilde{\Sigma}_0(\omega) \{1 - \exp[i\omega(t - t')]\} \\
 &= \int_{-\infty}^{\infty} \frac{d\omega}{\pi} \tilde{\Sigma}_0(\omega) \{1 - \cos[\omega(t - t')]\}. \quad (\text{G1})
 \end{aligned}$$

To bound the variance, rather than using the spectrum in the form (4.21), we use upper bounds. In the case for $1 < p < 3$, it is convenient to use the upper bound $\tilde{\Sigma}_0(\omega) < \kappa^{p-1}/|\omega|^p$, which gives, for $p \neq 2$,

$$\mathbb{E}[|X(t) - X(t')|^2] < (\kappa|t - t'|)^{p-1} \frac{-2\Gamma(1-p) \sin(\pi p/2)}{\pi}. \quad (\text{G2})$$

In the case $p = 2$, the result is $\kappa|t - t'|$, which is equivalent to taking the limit $p \rightarrow 2$ in Eq. (G2).

For $p \geq 3$, we upper bound the spectrum as

$$\tilde{\Sigma}_0(\omega) \leq \begin{cases} \kappa^{p-1}/\gamma^p & |\omega| < \gamma \\ \kappa^{p-1}/|\omega|^p & |\omega| \geq \gamma. \end{cases} \quad (\text{G3})$$

In the case $p = 3$, we then get

$$\begin{aligned}
 \mathbb{E}[|X(t) - X(t')|^2] &\leq \frac{(\kappa|t - t'|)^2}{\pi} \log\left(\frac{1}{\gamma|t - t'|}\right) \\
 &\quad + O[(\kappa|t - t'|)^2]. \quad (\text{G4})
 \end{aligned}$$

For $p > 3$,

$$\begin{aligned}
 \mathbb{E}[|X(t) - X(t')|^2] &\leq (\kappa|t - t'|)^2 \frac{p(\gamma/\kappa)^{3-p}}{3\pi(p-3)} \\
 &\quad + (\kappa|t - t'|)^{p-1} \frac{2\Gamma[2-p] \sin(\pi p/2)}{\pi(p-1)} \\
 &\quad + O[(\kappa|t - t'|)^3]. \quad (\text{G5})
 \end{aligned}$$

For $|t - t'| = T$, in each case, we find that the variance varies as a polynomial in κ/\mathcal{N} and is therefore small for large \mathcal{N}/κ .

APPENDIX H: MEASUREMENTS WITH BOUNDED ERROR

Here, we explain the modified measurement technique required in Sec. VII, to ensure that the probability of large phase errors is exponentially small in \mathcal{N}/κ . It is known that there are probe states, with maximum photon number $2N$ and mean photon number N , such that for a given integer M a canonical phase measurement accurately estimates any phase shift to within π/M , with an error probability exponentially small in N [71]. For each sample pulse, we can therefore apply such a measurement to a first subpulse having a number of photons scaling as $\Theta(\sqrt{\mathcal{N}T})$, with $M = 12$, providing an exponentially small probability of an error greater than $\pi/12$. Call this measurement A. For each sample pulse, a measurement as in Ref. [64] can also be performed on a second subpulse having a mean number of photons $\mathcal{N}T - \Theta(\sqrt{\mathcal{N}T})$. Call this measurement B. The mean-square error of this measurement is, to leading order, no different than the measurement with mean number $\mathcal{N}T$ because the difference in number is higher order.

The two measurements for each sample can be combined by using the result from measurement B if it is within $\pi/6$ of the result from measurement A. Provided the result from measurement A is no more than $\pi/12$ from the true phase, the result from measurement B will be no more than $\pi/4$ from the true phase. Hence, the probability of a phase error greater than $\pi/4$ is exponentially small. Conversely, if the two measurement results differ by more than $\pi/6$, then the result from measurement A is used. Provided the result from measurement A is no more than $\pi/12$ from the true phase, the error in the result from measurement A is then smaller than the error in measurement B. Hence, using the result from measurement A can only decrease the variance, while providing a probability of error greater than $\pi/4$ that is exponentially small. The net result of this combined measurement is that the probability of an error greater than $\pi/4$ can be made exponentially small in \mathcal{N}/κ , while not affecting the mean-square error to leading order.

APPENDIX I: PROOF OF FORMULA FOR a_m

To prove the last line of Eq. (7.16) for $m > 0$, we can simplify the sum over sinc functions to

$$\sum_{n=m}^{\infty} \text{sinc}(x - \pi n) = -\frac{(-1)^m \sin x}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{n + m - x/\pi}. \quad (\text{I1})$$

Using Eq. (5.7.7) of Ref. [72] gives

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n + z} = \frac{1}{2} \left[\Psi\left(\frac{1}{2} + \frac{z}{2}\right) - \Psi\left(\frac{z}{2}\right) \right]. \quad (\text{I2})$$

Combining the two, with $x = \pi t/T$ and $z = m - x/\pi$, gives the result in the last line of Eq. (7.16) for $m > 0$.

Similarly, for $m \leq 0$, we have

$$\sum_{n=-\infty}^{m-1} \text{sinc}(x - \pi n) = -\frac{(-1)^m \sin x}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{n + x/\pi + 1 - m}. \quad (\text{I3})$$

Combining this with Eq. (I2) gives the last line of Eq. (7.16) for $m \leq 0$.

APPENDIX J: TERMS IN EQ. (7.18)

Here, we justify the approximations used in the last line of Eq. (7.18) and show that the long-range correlations quantified by S are exponentially small.

1. The approximations

To justify the approximations, we first break up Z_n into contributions from the change in the phase and from the error in the samples. Using the definition of K_n in Eq. (7.4), we find that

$$Z_n = K_n - K_{n-1} = \frac{1}{2\pi} [\check{X}_n - \check{X}_{n-1} - X(nT) + X((n-1)T) - \xi_n + \xi_{n-1}]. \quad (\text{J1})$$

Now Z_n is an integer, and

$$\frac{1}{2\pi} |\check{X}_n - \check{X}_{n-1}| \leq 1/2. \quad (\text{J2})$$

This means that $[-X(nT) + X((n-1)T) - \xi_n + \xi_{n-1}]/(2\pi)$ takes values within $1/2$ of Z_n , which in turn implies that

$$Z_n = \text{round}\left(\frac{1}{2\pi} [-X(nT) + X((n-1)T) - \xi_n + \xi_{n-1}]\right). \quad (\text{J3})$$

Here, the rounding is taken to use the ‘‘round half up’’ convention. Now define

$$C_n := [-X(nT) + X((n-1)T)]/(2\pi), \quad (\text{J4})$$

$$D_n := [-\xi_n + \xi_{n-1}]/(2\pi), \quad (\text{J5})$$

so

$$Z_n = \text{round}(C_n + D_n). \quad (\text{J6})$$

The important thing to note is that, for $|n - m| > 1$, D_n is independent of D_m because these only depend on the error in independent measurements. Moreover, D_n and D_m are independent of C_n and C_m . However, C_n and C_m can be correlated because of correlations in the phase variation.

Using this notation and expanding $\mathbb{E}[Z_n Z_m]$ in terms of the probability distribution gives

$$\begin{aligned}
 \mathbb{E}[Z_n Z_m] &= \sum_{z_n} \sum_{z_m} \Pr(Z_n = z_n, Z_m = z_m) z_n z_m \\
 &= 2 \sum_{z_n > 0} \sum_{z_m > 0} \Pr(Z_n = z_n, Z_m = z_m) z_n z_m + 2 \sum_{z_n > 0} \sum_{z_m < 0} \Pr(Z_n = z_n, Z_m = z_m) z_n z_m \\
 &= 2 \sum_{z_n > 0} \sum_{z_m > 0} \Pr(Z_n \geq z_n, Z_m \geq z_m) - 2 \sum_{z_n > 0} \sum_{z_m < 0} \Pr(Z_n \geq z_n, Z_m \leq z_m) \\
 &= 2 \sum_{z_n > 0} \sum_{z_m > 0} [\Pr(C_n + D_n \geq z_n - 1/2, C_m + D_m \geq z_m - 1/2) \\
 &\quad - \Pr(C_n + D_n \geq z_n - 1/2, C_m + D_m \leq -z_m + 1/2)]. \tag{J7}
 \end{aligned}$$

Considering this expression for $n = m$, it simplifies to

$$\mathbb{E}[Z_m^2] = 2 \sum_{z_m > 0} \Pr(C_m + D_m \geq z_m - 1/2). \tag{J8}$$

This is because the probability on the last line of (J7) is equal to zero since the two conditions $C_n + D_n \geq z_n - 1/2$ and $C_m + D_m \leq -z_m + 1/2$ are incompatible (except in the case of equality, which has zero probability). Now it is important to note that C_m is a Gaussian random variable with a variance that decreases as \mathcal{N} is increased. It can be immediately inferred that it is Gaussian because it is a linear combination of $X(mT)$ and $X((m-1)T)$, both of which are Gaussian. It has zero mean because the expectation values of $X(mT)$ and $X((m-1)T)$ are equal. Its variance is given by

$$\langle C_m^2 \rangle = \langle [X(mT) - X((m-1)T)]^2 \rangle / (2\pi)^2. \tag{J9}$$

Using the result in Appendix G, this variance decreases polynomially in T as T is decreased. Because we choose

T to decrease polynomially as \mathcal{N} is increased, so does the variance of C_m .

In addition, $|D_m|$ can be no larger than 1, and the phase measurements may be chosen such that the probability of $|D_m|$ larger than $1/4$ is exponentially small in \mathcal{N} . As a result, the terms in the sum in Eq. (J8) decrease exponentially in the product of \mathcal{N} and z_m . Hence, the terms for $z_m > 1$ are exponentially small in \mathcal{N} as compared to that for $z_m = 1$, and we have

$$\mathbb{E}[Z_m^2] \approx 2 \Pr(C_m + D_m \geq 1/2) = \Pr(|Z_m| > 0) = p_{\text{err}}. \tag{J10}$$

The other approximation in the last line of Eq. (7.18) follows immediately because $\mathbb{E}[Z_m^2] \geq |\mathbb{E}[Z_n Z_m]|$, so $\mathbb{E}[Z_m Z_{m+1}] \lesssim p_{\text{err}}$.

2. Bounding S

Lastly, we wish to show that S , defined in Eq. (7.19), is exponentially small in \mathcal{N} . For $|n - m| > 1$, we evaluate

$$\begin{aligned}
 &\Pr(C_n + D_n \geq z_n - 1/2, C_m + D_m \geq z_m - 1/2) - \Pr(C_n + D_n \geq z_n - 1/2, C_m + D_m \leq -z_m + 1/2) \\
 &= \int dd_n \int dd_m [\Pr(C_n \geq z_n - 1/2 - d_n, C_m \geq z_m - 1/2 - d_m) \\
 &\quad - \Pr(C_n \geq z_n - 1/2 - d_n, C_m \leq -z_m + 1/2 + d_m)] P(D_n = d_n) P(D_m = d_m). \tag{J11}
 \end{aligned}$$

In Eq. (J11), we denote the probability density by P , and in the last line, we have used the symmetry of the probability density $P(D_m = d_m)$ about zero.

It is convenient to denote

$$v_m := z_m - 1/2 - d_m, \quad v_n := z_n - 1/2 - d_n. \tag{J12}$$

Let us denote $\sigma^2 := \mathbb{E}[C_n^2]$ and $V_{n-m} := \mathbb{E}[C_n C_m]$. Then, the probability density is

$$P(C_n = x, C_m = y) = \frac{1}{2\pi \sqrt{\sigma^4 - V_{n-m}^2}} \exp \left[-\frac{\sigma^2(x^2 + y^2) - 2V_{n-m}xy}{2(\sigma^4 - V_{n-m}^2)} \right]. \tag{J13}$$

Let us use the notation $V := |V_{n-m}|$ and $\varsigma := \text{sign}(V_{n-m})$. Note that these quantities depend on $n - m$; we do not show this dependence, so the equations may be presented more simply. Then, the probability density can be given as

$$P(C_n = x, C_m = y) = \frac{1}{2\pi\sqrt{\sigma^4 - V^2}} \exp \left[-\frac{x^2 + y^2}{2(\sigma^2 + V)} - \frac{V(x - \varsigma y)^2}{2(\sigma^4 - V^2)} \right]. \quad (\text{J14})$$

In terms of this probability distribution, the difference of probabilities is

$$\begin{aligned} & \Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m) \\ &= \int_{v_n}^{\infty} dx \int_{v_m}^{\infty} dy [P(C_n = x, C_m = y) - P(C_n = x, C_m = -y)] \\ &= \frac{1}{2\pi\sqrt{\sigma^4 - V^2}} \int_{v_n}^{\infty} dx \int_{v_m}^{\infty} dy \exp \left[-\frac{x^2 + y^2}{2(\sigma^2 + V)} \right] \left\{ \exp \left[-\frac{V(x - \varsigma y)^2}{2(\sigma^4 - V^2)} \right] - \exp \left[-\frac{V(x + \varsigma y)^2}{2(\sigma^4 - V^2)} \right] \right\} \\ &= \frac{1}{2\pi\sqrt{\sigma^4 - V^2}} \int_{|v_n|}^{\infty} dx \int_{|v_m|}^{\infty} dy \exp \left[-\frac{x^2 + y^2}{2(\sigma^2 + V)} \right] \left\{ \exp \left[-\frac{V(x - \varsigma y)^2}{2(\sigma^4 - V^2)} \right] - \exp \left[-\frac{V(x + \varsigma y)^2}{2(\sigma^4 - V^2)} \right] \right\}. \end{aligned} \quad (\text{J15})$$

In the last line, we used the fact that, if v_m is negative, the components of the integral for negative values of y cancel out the components of the integral for the corresponding positive values of y . Similarly, if v_n is negative, the components for negative x cancel with those for corresponding positive values of x . Since the argument of the exponential is negative, its derivative is no greater than 1, and we have

$$|\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| \leq \frac{1}{2\pi\sqrt{\sigma^4 - V^2}} \int_{|v_n|}^{\infty} dx \int_{|v_m|}^{\infty} dy \exp \left[-\frac{x^2 + y^2}{2(\sigma^2 + V)} \right] \frac{2Vxy}{\sigma^4 - V^2}. \quad (\text{J16})$$

This result then gives

$$\begin{aligned} |\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| &\leq \frac{V}{\pi(\sigma^4 - V^2)^{3/2}} \int_{|v_n|}^{\infty} dx \int_{|v_m|}^{\infty} dy xy \exp \left[-\frac{x^2 + y^2}{2(\sigma^2 + V)} \right] \\ &= \frac{V\sqrt{\sigma^2 + V}}{\pi(\sigma^2 - V)^{3/2}} \exp \left[-\frac{v_n^2 + v_m^2}{2(\sigma^2 + V)} \right]. \end{aligned} \quad (\text{J17})$$

An alternative bound can be derived from Eq. (J15). If $V_{n-m} > 0$, then the first term in braces in Eq. (J15) is larger, and if $V_{n-m} < 0$, then the second term in braces in Eq. (J15) is larger. In either case, we find

$$\begin{aligned} & |\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| \\ &\leq \frac{1}{2\pi\sqrt{\sigma^4 - V^2}} \int_{|v_n|}^{\infty} dx \int_{|v_m|}^{\infty} dy \exp \left[-\frac{x^2 + y^2}{2(\sigma^2 + V)} \right] \exp \left[-\frac{V(x - y)^2}{2(\sigma^4 - V^2)} \right] \\ &= \frac{1}{2\pi\sqrt{\sigma^4 - V^2}} \int_{|v_n|}^{\infty} dx \int_{|v_m|}^{\infty} dy \exp \left[-\frac{(x + y)^2}{4(\sigma^2 + V)} \right] \exp \left[-\frac{(x - y)^2}{4(\sigma^2 - V)} \right] \\ &= \frac{1}{4\pi\sqrt{\sigma^4 - V^2}} \int_{|v_n|+|v_m|}^{\infty} du \int_{2|v_n|-u}^{u-2|v_m|} dw \exp \left[-\frac{u^2}{4(\sigma^2 + V)} \right] \exp \left[-\frac{w^2}{4(\sigma^2 - V)} \right], \end{aligned} \quad (\text{J18})$$

where we have made the change of variables $u = x + y$ and $w = x - y$. Because the integral over w is upper bounded by the integral over the infinite interval, we have

$$\begin{aligned} |\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| &\leq \frac{1}{2\sqrt{\pi(\sigma^2 + V)}} \int_{|v_n|+|v_m|}^{\infty} du \exp \left[-\frac{u^2}{4(\sigma^2 + V)} \right] \\ &= \frac{1}{2} \text{erfc} \left[\frac{|v_n| + |v_m|}{2\sqrt{\sigma^2 + V}} \right] \leq \frac{1}{2} \exp \left[-\frac{(|v_n| + |v_m|)^2}{4(\sigma^2 + V)} \right]. \end{aligned} \quad (\text{J19})$$

Because $V \leq \sigma^2$ and $(|v_n| + |v_m|)^2 \geq v_n^2 + v_m^2$, we can use this to obtain the inequality

$$|\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| \leq \frac{1}{2} \exp \left[-\frac{v_n^2 + v_m^2}{8\sigma^2} \right]. \quad (\text{J20})$$

We can rewrite Eq. (J17) as

$$|\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| \leq \frac{(V/\sigma^2)\sqrt{1+V/\sigma^2}}{\pi(1-V/\sigma^2)^{3/2}} \exp \left[-\frac{v_n^2 + v_m^2}{2(\sigma^2 + V)} \right]. \quad (\text{J21})$$

Again using $V \leq \sigma^2$, this implies

$$|\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| \leq \frac{(V/\sigma^2)\sqrt{1+V/\sigma^2}}{\pi(1-V/\sigma^2)^{3/2}} \exp \left[-\frac{v_n^2 + v_m^2}{8\sigma^2} \right]. \quad (\text{J22})$$

Combining this result with Eq. (J20) gives

$$|\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| \leq \min \left\{ \frac{1}{2}, \frac{(V/\sigma^2)\sqrt{1+V/\sigma^2}}{\pi(1-V/\sigma^2)^{3/2}} \right\} \exp \left[-\frac{v_n^2 + v_m^2}{8\sigma^2} \right]. \quad (\text{J23})$$

The possible range of V/σ^2 is $[0,1]$. It is easily checked that, within this range, the minimum in the above expression is no larger than $1.04 \times V/\sigma^2$. Hence, we have the inequality

$$|\Pr(C_n \geq v_n, C_m \geq v_m) - \Pr(C_n \geq v_n, C_m \leq -v_m)| \leq 1.04 \frac{|V_{n-m}|}{\sigma^2} \exp \left[-\frac{v_n^2 + v_m^2}{8\sigma^2} \right]. \quad (\text{J24})$$

It is important to note that σ^2 decreases polynomially in \mathcal{N} . Hence, for $z_n > 1$ or $z_m > 1$, this expression is exponentially small in \mathcal{N} . Moreover, this expression decays exponentially in z_n and z_m , so the sum over $z_n > 0$ and $z_m > 0$, excluding $z_n = z_m = 1$, is exponentially small in \mathcal{N} . Considering the remaining term for $z_n = z_m = 1$, we obtain the bound

$$\begin{aligned} & \int dd_n \int dd_m \exp \left(-\frac{(1/2 - d_n)^2 + (1/2 - d_m)^2}{8\sigma^2} \right) \\ & \quad \times P(D_n = d_n)P(D_m = d_m) \\ & = \left[\int dd_n \exp \left(-\frac{(1/2 - d_n)^2}{8\sigma^2} \right) P(D_n = d_n) \right]^2 \\ & \leq \left[\Pr(D_n > 1/4) + \exp \left(-\frac{1}{128\sigma^2} \right) \right]^2. \end{aligned} \quad (\text{J25})$$

It is possible to perform measurements such that the probability of error larger than $\pi/4$ is exponentially small in \mathcal{N} . As a result, using Eq. (J7), we have

$$\mathbb{E}[Z_n Z_m] = |V_{n-m}| g(\mathcal{N}), \quad (\text{J26})$$

where $g(\mathcal{N})$ is an exponentially decreasing function.

Evaluating V_{n-m} , we obtain

$$\begin{aligned} V_{n-m} & = \mathbb{E}[C_n C_m] = \mathbb{E}[\{X((n-1)T) - X(nT)\} \\ & \quad \times \{X((m-1)T) - X(mT)\}] \\ & = 2\Sigma_0((n-m)T) - \Sigma_0((n-m+1)T) \\ & \quad - \Sigma_0((n-m-1)T) \\ & = \frac{1}{\pi} \int d\omega \tilde{\Sigma}_0(\omega) \cos[\omega(n-m)T] [1 - \cos \omega T]. \end{aligned} \quad (\text{J27})$$

Taking $m = 0$ and $n > 0$ for simplicity, and integrating by parts, we obtain

$$\begin{aligned} V_n & = -\frac{1}{\pi} \int_0^\infty d\omega \tilde{\Sigma}'_0(\omega) \left\{ \frac{1}{nT} \sin(\omega n T) - \frac{1}{2(n-1)T} \right. \\ & \quad \left. \times \sin[\omega(n-1)T] - \frac{1}{2(n+1)T} \sin[\omega(n-1)T] \right\}, \end{aligned} \quad (\text{J28})$$

where we have used the fact that $\tilde{\Sigma}_0(\omega)$ is bounded at $\omega = 0$ and approaches zero for $\omega \rightarrow \infty$. Using the fact that $\tilde{\Sigma}'_0(\omega) \leq 0$,

$$\begin{aligned}
|V_n| &\leq -\frac{1}{\pi} \int_0^\infty d\omega \tilde{\Sigma}'_0(\omega) \left\{ \frac{1}{nT} + \frac{1}{2(n-1)T} + \frac{1}{2(n+1)T} \right\} \\
&= \frac{1}{\pi} \tilde{\Sigma}_0(0) \left\{ \frac{1}{nT} + \frac{1}{2(n-1)T} + \frac{1}{2(n+1)T} \right\} \\
&\leq \frac{7}{3\pi} \tilde{\Sigma}_0(0) \frac{1}{Tn}.
\end{aligned} \tag{J29}$$

As $\tilde{\Sigma}_0(0)$ is bounded (equal to κ^{p-1}/γ^p), V_{n-m} scales as $1/|n-m|$. Hence, $\mathbb{E}[Z_n Z_m]$ scales as a factor exponentially small in \mathcal{N} times $1/|n-m|$.

Next, we consider the scaling for a_m . Using the first two terms of the asymptotic series for the digamma function,

$$\Psi(x) = \ln(x) - \frac{1}{2x} + O(x^{-2}), \tag{J30}$$

we have

$$\begin{aligned}
\Psi(x+1/2) - \Psi(x) &= \ln(x+1/2) - \ln(x) - \frac{1}{2x+1} + \frac{1}{2x} + O(x^{-2}) \\
&= \frac{1}{2x} + O(x^{-2}).
\end{aligned} \tag{J31}$$

As a result, a_m has the scaling

$$a_m = -(-1)^m \sin\left(\frac{\pi t}{T}\right) \frac{1}{|m|} + O(m^{-2}). \tag{J32}$$

In fact, it turns out that $|a_m| \leq \pi/|m|$.

Now we have sufficient results to bound the component of the sum

$$S = \sum_{n,m:|n-m|>1} \mathbb{E}[Z_n Z_m] a_m a_n. \tag{J33}$$

Using the above results, the sum can be bounded as

$$S \lesssim f(\mathcal{N}) \sum_{n,m:|n-m|>1, n \neq 0, m \neq 0} \frac{1}{|n-m| \times |n| \times |m|}, \tag{J34}$$

where f is an exponentially decreasing function. Splitting the sum into $m < n$ and $m > n$, the bound can be rewritten as

$$S \lesssim 4f(\mathcal{N}) \sum_{m>n>1} \frac{1}{|n-m|nm} = 4f(\mathcal{N}) \sum_{n,r>0} \frac{1}{n(n+r)r}. \tag{J35}$$

Using the inequality $x^2 + y^2 \geq 2xy$ for $x = \sqrt{n}$ and $y = \sqrt{r}$ gives $n+r \geq 2\sqrt{nr}$. Hence, substituting $n+r$ with $2\sqrt{nr}$ gives

$$\begin{aligned}
S &\lesssim 4f(\mathcal{N}) \sum_{n,r>0} \frac{1}{n(2\sqrt{nr})r} \\
&= 2f(\mathcal{N}) \sum_{n,r>0} \frac{1}{n^{3/2}r^{3/2}} \\
&= 2f(\mathcal{N}) [\zeta(3/2)]^2 \\
&\leq 14f(\mathcal{N}),
\end{aligned} \tag{J36}$$

where $\zeta(z)$ is the Riemann zeta function. This result means that the sum S is exponentially small in \mathcal{N} , as claimed.

-
- [1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
 - [2] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum-Enhanced Measurements: Beating the Standard Quantum Limit*, *Science* **306**, 1330 (2004).
 - [3] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, England, 2010).
 - [4] V. Giovannetti, S. Lloyd, and L. Maccone, *Advances in Quantum Metrology*, *Nat. Photonics* **5**, 222 (2011).
 - [5] A. S. Holevo, *Quantum Systems, Channels, Information* (de Gruyter, Berlin, 2012).
 - [6] V. B. Braginsky and F. Ya. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, England, 1992).
 - [7] M. Tsang, H. M. Wiseman, and C. M. Caves, *Fundamental Quantum Limit to Waveform Estimation*, *Phys. Rev. Lett.* **106**, 090401 (2011).
 - [8] M. Tsang and R. Nair, *Fundamental Quantum Limits to Waveform Detection*, *Phys. Rev. A* **86**, 042115 (2012).
 - [9] M. Tsang, *Quantum Metrology with Open Dynamical Systems*, *New J. Phys.* **15**, 073005 (2013).
 - [10] N. Treps, N. Grosse, W. P. Bowen, C. Fabre, H.-A. Bachor, and P. K. Lam, *A Quantum Laser Pointer*, *Science* **301**, 940 (2003).
 - [11] M. Tsang, *Quantum Imaging Beyond the Diffraction Limit by Optical Centroid Measurements*, *Phys. Rev. Lett.* **102**, 253601 (2009).
 - [12] M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, and W. P. Bowen, *Biological Measurement Beyond the Quantum Limit*, *Nat. Photonics* **7**, 229 (2013).
 - [13] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, *Optimal Frequency Measurements with Maximally Correlated States*, *Phys. Rev. A* **54**, R4649 (1996).
 - [14] B. Yurke, S. L. McCall, and J. R. Klauder, *SU(2) and SU(1,1) Interferometers*, *Phys. Rev. A* **33**, 4033 (1986).
 - [15] B. C. Sanders and G. J. Milburn, *Optimal Quantum Measurements for Phase Estimation*, *Phys. Rev. Lett.* **75**, 2944 (1995).
 - [16] Z. Y. Ou, *Complementarity and Fundamental Limit in Precision Phase Measurement*, *Phys. Rev. Lett.* **77**, 2352 (1996).
 - [17] M. Zwiernik, C. A. Pérez-Delgado, and P. Kok, *General Optimality of the Heisenberg Limit for Quantum Metrology*, *Phys. Rev. Lett.* **105**, 180402 (2010); **107**, 059904(E) (2011).

- [18] Á. Rivas and A. Luis, *Sub-Heisenberg Estimation of Non-random Phase Shifts*, *New J. Phys.* **14**, 093052 (2012).
- [19] A. Luis and A. Rodil, *Alternative Measures of Uncertainty in Quantum Metrology: Contradictions and Limits*, *Phys. Rev. A* **87**, 034101 (2013).
- [20] A. Luis, *Signal Detection without Finite-Energy Limits to Quantum Resolution*, *Ann. Phys. (Amsterdam)* **331**, 1 (2013).
- [21] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling, *Quantum Metrology with Two-Mode Squeezed Vacuum: Parity Detection Beats the Heisenberg Limit*, *Phys. Rev. Lett.* **104**, 103602 (2010).
- [22] Y. R. Zhang, G. R. Jin, J. P. Cao, W. M. Liu, and H. Fan, *Unbounded Quantum Fisher Information in Two-Path Interferometry with Finite Photon Number*, *J. Phys. A* **46**, 035302 (2013).
- [23] M. Tsang, *Ziv-Zakai Error Bounds for Quantum Parameter Estimation*, *Phys. Rev. Lett.* **108**, 230401 (2012).
- [24] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum Measurement Bounds Beyond the Uncertainty Relations*, *Phys. Rev. Lett.* **108**, 260405 (2012).
- [25] D. W. Berry, M. J. W. Hall, M. Zwierz, and H. M. Wiseman, *Optimal Heisenberg-Style Bounds for the Average Performance of Arbitrary Phase Estimates*, *Phys. Rev. A* **86**, 053813 (2012).
- [26] M. J. W. Hall, D. W. Berry, M. Zwierz, and H. M. Wiseman, *Universality of the Heisenberg Limit for Estimates of Random Phase Shifts*, *Phys. Rev. A* **85**, 041802 (2012).
- [27] R. Nair, *Fundamental Limits on the Accuracy of Optical Phase Estimation from Rate-Distortion Theory*, arXiv:1204.3761.
- [28] V. Giovannetti and L. Maccone, *Sub-Heisenberg Estimation Strategies are Ineffective*, *Phys. Rev. Lett.* **108**, 210404 (2012).
- [29] M. J. W. Hall and H. M. Wiseman, *Does Nonlinear Metrology Offer Improved Resolution? Answers from Quantum Information Theory*, *Phys. Rev. X* **2**, 041006 (2012).
- [30] S. Knysh, V. N. Smelyanskiy, and G. A. Durkin, *Scaling Laws for Precision in Quantum Interferometry and the Bifurcation Landscape of the Optimal State*, *Phys. Rev. A* **83**, 021804 (2011).
- [31] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, *General Framework for Estimating the Ultimate Precision Limit in Noisy Quantum-Enhanced Metrology*, *Nat. Phys.* **7**, 406 (2011).
- [32] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, *Quantum Metrology for Noisy Systems*, *Braz. J. Phys.* **41**, 229 (2011).
- [33] B. M. Escher, L. Davidovich, N. Zagury, and R. L. de Matos Filho, *Quantum Metrological Limits via a Variational Approach*, *Phys. Rev. Lett.* **109**, 190404 (2012).
- [34] C. L. Latune, B. M. Escher, R. L. de Matos Filho, and L. Davidovich, *Quantum Limit for the Measurement of a Classical Force Coupled to a Noisy Quantum-Mechanical Oscillator*, *Phys. Rev. A* **88**, 042112 (2013).
- [35] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, *The Elusive Heisenberg Limit in Quantum-Enhanced Metrology*, *Nat. Commun.* **3**, 1063 (2012).
- [36] S. I. Knysh, E. H. Chen, and G. A. Durkin, *True Limits to Precision via Unique Quantum Probe*, arXiv:1402.0495.
- [37] M. I. Kolobov, *The Spatial Behavior of Nonclassical Light*, *Rev. Mod. Phys.* **71**, 1539 (1999).
- [38] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, *Quantum Enhanced Multiple Phase Estimation*, *Phys. Rev. Lett.* **111**, 070403 (2013).
- [39] *Quantum State Estimation*, edited by M. G. A. Paris and J. Řeháček (Springer-Verlag, Berlin, 2004).
- [40] K. C. Young, M. Sarovar, R. Kosut, and K. B. Whaley, *Optimal Quantum Multiparameter Estimation and Application to Dipole- and Exchange-Coupled Qubits*, *Phys. Rev. A* **79**, 062301 (2009).
- [41] D. W. Berry, M. J. W. Hall, and H. M. Wiseman, *Stochastic Heisenberg Limit: Optimal Estimation of a Fluctuating Phase*, *Phys. Rev. Lett.* **111**, 113601 (2013).
- [42] 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, *Adaptive Optical Phase Estimation Using Time-Symmetric Quantum Smoothing*, *Phys. Rev. Lett.* **104**, 093601 (2010).
- [43] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph, H. M. Wiseman, E. H. Huntington, and A. Furusawa, *Quantum-Enhanced Optical-Phase Tracking*, *Science* **337**, 1514 (2012).
- [44] K. Iwasawa, K. Makino, H. Yonezawa, M. Tsang, A. Davidovic, E. Huntington, and A. Furusawa, *Quantum-Limited Mirror-Motion Estimation*, *Phys. Rev. Lett.* **111**, 163602 (2013).
- [45] H. P. Yuen and M. Lax, *Multiple-Parameter Quantum Estimation and Measurement of Nonselfadjoint Observables*, *IEEE Trans. Inf. Theory* **19**, 740 (1973).
- [46] C. W. Helstrom and R. S. Kennedy, *Noncommuting Observables in Quantum Detection and Estimation Theory*, *IEEE Trans. Inf. Theory* **20**, 16 (1974).
- [47] M. G. A. Paris, *Quantum Estimation for Quantum Technology*, *Int. J. Quantum. Inform.* **07**, 125 (2009).
- [48] The quantum Cramér-Rao bound gives a lower bound on the mean-square phase-estimation error of $1/(\Delta n)^2$, where Δn^2 is photon-number variance. This is far weaker than the Heisenberg limit of $1/\bar{n}^2$, as can be seen from the state $\frac{\sqrt{3}}{2} \sum_{n=0}^{\infty} 2^{-n} |2^n\rangle$ which has $\bar{n} = 3/2$ but divergent Δn .
- [49] Y.-R. Zhang and H. Fan, *Quantum Metrological Bounds for Vector Parameters*, *Phys. Rev. A* **90**, 043818 (2014).
- [50] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory* (John Wiley & Sons, New York, 2001).
- [51] The terminology “Bell-Ziv-Zakai bounds” was adopted in Ref. [56].
- [52] K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees, *Extended Ziv-Zakai Lower Bound for Vector Parameter Estimation*, *IEEE Trans. Inf. Theory* **43**, 624 (1997).
- [53] J. Ziv and M. Zakai, *Some Lower Bounds on Signal Parameter Estimation*, *IEEE Trans. Inf. Theory* **15**, 386 (1969).
- [54] S. Basu and Y. Bresler, *A Global Lower Bound on Parameter Estimation Error with Periodic Distortion Functions*, *IEEE Trans. Inf. Theory* **46**, 1145 (2000).


- [55] D. W. Berry and H. M. Wiseman, *Adaptive Phase Measurements for Narrowband Squeezed Beams*, *Phys. Rev. A* **73**, 063824 (2006); **87**, 019901(E) (2013).
- [56] *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering Tracking*, edited by H. L. Van Trees and K. L. Bell (Wiley-IEEE, Piscataway, 2007).
- [57] T. Kailath, *The Divergence and Bhattacharyya Distance Measures in Signal Selection*, *IEEE Transactions on Communication Technology* **15**, 52 (1967).
- [58] G. T. Toussaint, *Comments on "The Divergence and Bhattacharyya Distance Measures in Signal Selection"*, *IEEE Trans. Commun. Tech.* **20**, 485 (1972).
- [59] A. Uhlmann, *The "Transition Probability" in the State Space of a *-Algebra*, *Rep. Math. Phys.* **9**, 273 (1976).
- [60] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2011).
- [61] M. J. W. Hall, *Gaussian Noise and Quantum-Optical Communication*, *Phys. Rev. A* **50**, 3295 (1994).
- [62] R. Loudon, *The Quantum Theory of Light* (Oxford University Press, New York, 2000).
- [63] J. H. Shapiro, *Quantum Measurement Eigenkets for Continuous-Time Direct Detection*, *Quantum and Semiclassical Optics: J. Eur. Opt. Soc. Part B* **10**, 567 (1998).
- [64] A. Bandilla, H. Paul, and H.-H. Ritze, *Realistic Quantum States of Light with Minimum Phase Uncertainty*, *Quantum Optics: J. Eur. Opt. Soc. Part B* **3**, 267 (1991).
- [65] H. M. Wiseman, *Adaptive Phase Measurements of Optical Modes: Going Beyond the Marginal Q Distribution*, *Phys. Rev. Lett.* **75**, 4587 (1995).
- [66] D. W. Berry and H. M. Wiseman, *Adaptive Quantum Measurements of a Continuously Varying Phase*, *Phys. Rev. A* **65**, 043803 (2002).
- [67] M. Tsang, J. H. Shapiro, and S. Lloyd, *Quantum Theory of Optical Temporal Phase and Instantaneous Frequency*, *Phys. Rev. A* **78**, 053820 (2008).
- [68] M. Tsang, J. H. Shapiro, and S. Lloyd, *Quantum Theory of Optical Temporal Phase and Instantaneous Frequency. II. Continuous-Time Limit and State-Variable Approach to Phase-Locked Loop Design*, *Phys. Rev. A* **79**, 053843 (2009).
- [69] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part II: Nonlinear Modulation Theory* (John Wiley & Sons, New York, 2003).
- [70] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, *Adaptive Homodyne Measurement of Optical Phase*, *Phys. Rev. Lett.* **89**, 133602 (2002).
- [71] M. J. W. Hall and I. G. Fuss, *Quantum Phase Detection and Digital Communication*, *Quantum Optics: J. Eur. Opt. Soc. Part B* **3**, 147 (1991).
- [72] *NIST Handbook of Mathematical Functions*, edited by F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (Cambridge University Press, New York, NY, 2010).

[Log in to My Ulrich's](#)

Macquarie University Library --Select Language--

[Search](#) [Workspace](#) [Ulrich's Update](#) [Admin](#)

Enter a Title, ISSN, or search term to find journals or other periodicals:

2160-3308 

[▶ Advanced Search](#)



Search My Library's Catalog: [ISSN Search](#) | [Title Search](#)

[Search Results](#)

Physical Review X

Title Details

Lists



[Marked Titles \(0\)](#)

Search History

[2160-3308 - \(1\)](#)

 Save to List  Email  Download  Print  Corrections  Expand All  Collapse All

▼ Basic Description

Title	Physical Review X
ISSN	2160-3308
Publisher	American Physical Society
Country	United States
Status	Active
Start Year	2011
Frequency	Irregular
Earliest Volume Note	Fall
Language of Text	Text in: English
Refereed 	Yes
Abstracted / Indexed	Yes
Open Access 	Yes http://prx.aps.org
Serial Type	Journal
Content Type	Academic / Scholarly
Format	Online
Website	http://prx.aps.org
Description	Provides open access, primary research covering all of physics and its application to related fields.

▶ Subject Classifications

▶ Additional Title Details

▶ Publisher & Ordering Details

▶ Online Availability

▶ Abstracting & Indexing

▶ Demographics

▶ Reviews

 Save to List  Email  Download  Print  Corrections  Expand All  Collapse All