Stochastic Waveform Estimation at the Fundamental Quantum Limit

James W. Gardner,[1,2,*] Tuvia Gefen,[3,†] Simon A. Haine,[4] Joseph J. Hope,[4] John Preskill,[2,3] Yanbei Chen,[2] and Lee McCuller[3,5]

[1] *OzGrav-ANU, Centre for Gravitational Astrophysics, Research Schools of Physics, and of Astronomy and Astrophysics, The Australian National University, Canberra, Australian Capital Territory 2601, Australia*

[2] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA*

[3] *Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

[4] *Department of Quantum Science and Technology and Department of Fundamental and Theoretical Physics, Research School of Physics, The Australian National University, Canberra, Australian Capital Territory 0200, Australia*

[5] *Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, California 91125, USA*

(Received 1 May 2024; revised 9 May 2025; accepted 12 June 2025; published 22 July 2025)

Although measuring the deterministic waveform of a weak classical force is a well-studied problem, estimating a random waveform, such as the spectral density of a stochastic signal field, is much less well understood despite it being a widespread task at the frontier of experimental physics. State-of-the-art precision sensors of random forces must account for the underlying quantum nature of the measurement but the optimal quantum protocol for interrogating such linear sensors is not known. We derive the fundamental precision limit: the extended-channel quantum Cramér-Rao bound. In the experimentally relevant regime in which losses dominate, we prove that non-Gaussian-state preparation and measurement are required to achieve this fundamental limit and we determine numerically the optimal non-Gaussian protocol. We discuss how this scheme could accelerate searches for signatures of quantum gravity, stochastic gravitational waves, and axionic dark matter.

DOI: 10.1103/h91r-4ws9

I. INTRODUCTION

The estimation of the spectral density of a classical process is a ubiquitous task in continuous-variable quantum systems. Examples include searching for excess noise in optical interferometers due to quantum gravity [1–5], probing stochastic gravitational waves with the global network of gravitational-wave observatories [6–10], and hunting for axionic dark matter with microwave cavities [11–14]. The remarkable sensitivity of contemporary devices demands that we contend with quantum noise, the fundamental uncertainty in the state of the device arising from the Heisenberg uncertainty principle. The stochastic signal of interest must be distinguished from the natural fluctuations arising from the measurement of the device.

This task, sometimes called "noise spectroscopy," can be expressed in the language of quantum metrology, the study of estimating parameters encoded in quantum states. We consider a quantum device coupled to a stochastic signal, a classical continuous random variable $y(t)$ at times t. We assume this signal to be a Gaussian stationary random process; assuming that the signal is Gaussian is valid provided that the central limit theorem applies, e.g., that we repeat the experiment many times. Assuming that the signal is a stationary random process allows us to study its power spectral density $S_{yy}(\Omega)$ at each positive frequency Ω. Our goal is to estimate the continuum values of $S_{yy}(\Omega)$ by measuring the state of the quantum device. We want our estimate at each Ω to be unbiased and have the minimal mean-square error (MSE). We will show that this resembles the problem of estimating the uncertainty of a single bosonic mode and so we treat that case extensively.

Much attention has previously been dedicated to studying the fundamental precision limits of estimating the

*Contact author: james.gardner@anu.edu.au
†Contact author: tgefen@caltech.edu

mean values of the signal for various protocols and quantum devices (see, e.g., Refs. [15–18]). Here, we focus instead on estimating the spectral density $S_{yy}(\Omega)$ of the signal rather than its mean value. This is a much less studied problem. It has recently been shown that this scenario is fundamentally different from mean-value estimation in that non-Gaussian measurement techniques are required to obtain optimal estimates [19]. There are still many critical open questions, however, such as the effects of imperfections and the preparation of different initial quantum states.

To guide the reader, we summarize the main results of our work as follows:

(1) In Sec. II, we reduce the noise-estimation problem for a linear quantum device at a fixed frequency Ω to the study of a harmonic oscillator undergoing a random-displacement channel. The task is to estimate the standard deviation σ of the Gaussian distribution of random displacements and, in some cases, to simultaneously estimate the mean displacement μ.

(2) In Sec. III, we review the prior literature on estimating σ in the ideal lossless limit, using an initial vacuum state. A number-resolving measurement, e.g., photon counting, is optimal [19,20]. In contrast, Gaussian measurements, such as homodyne detection, suffer the "Rayleigh curse" when the signal σ is small. The goal of our work is to find the optimal initial state and measurement scheme that achieves the ultimate limit on precision, the extended-channel quantum Fisher information (ECQFI), in the presence of loss. We also consider the simultaneous estimation of μ and σ.

(3) In Sec. IV, we discuss imperfections. We focus on the case of a loss η occurring before the signal is encoded by the random-displacement channel. We assume negligible loss after the encoding and negligible additive classical noise, σ_C. Experimentally, the relevant regime is that of small signals $\sigma^2 \ll 1/2$ and high loss $\eta \gg \sigma^2$.

(4) In Sec. V, we find the optimal protocol for estimating σ. In the lossless case, in which the initial state obeys the energy constraint $\langle \hat{n} \rangle = N$, we show that the ECQFI is saturated by preparing an initial single-mode squeezed vacuum (SMSV) and then, after the encoding, antisqueezing and performing a number measurement. In the lossy case, we show that preparing a two-mode squeezed vacuum (TMSV) state is optimal but only if negligible loss occurs on the ancilla mode. In the experimentally relevant regime of high loss on all modes, numerics indicate that the ECQFI is attainable, without an ancilla, using highly non-Gaussian states and measurements.

(5) In Sec. VI, we discuss how to simultaneously estimate μ and σ. We assume that the initial state is vacuum (which may not be optimal). For separable measurements on M copies of the final state, we construct an adaptive measurement scheme in which we learn about μ via quadrature measurement and then use that information to displace back to the origin and learn about σ via number measurement. To saturate the quantum Fisher information (QFI) for the fixed-vacuum input state, however, a collective measurement on the M modes is required.

(6) Finally, in Sec. VII, we discuss how our results concerning single-mode channel estimation can be leveraged for estimating the power spectral density of a continuously varying signal, we propose experimental implementations to realize the optimal estimation protocols, and we apply our results to searches for signatures of quantum gravity, stochastic gravitational waves, and axionic dark matter.

II. NOISE ESTIMATION OF A LINEAR QUANTUM DEVICE

We now consider estimating $S_{yy}(\Omega)$ at each frequency Ω from measurements of the quantum device. For a linear device, we show that this is equivalent to estimating the random displacement of a harmonic oscillator.

A. Review of linear quantum devices

We consider a quantum device that responds linearly to the continuous stochastic signal $y(t)$ and is coupled to an environment of incoming and outgoing bosonic modes as shown in Fig. 1(a). This is a valid model for many relevant quantum systems. For example, the device could be a gravitational-wave interferometer coupled to a stochastic gravitational-wave signal and to the light entering and exiting the differential port of the Michelson interferometer. In the input-output formalism [21], let the annihilation operator of the outgoing bosonic mode be $\hat{a}(t)$,

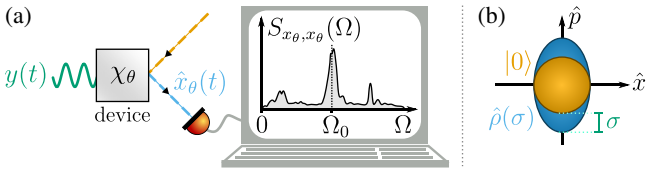

FIG. 1. (a) The spectral density of the outgoing bosonic mode from a linear quantum device coupled to a classical random process $y(t)$, e.g., a gravitational-wave observatory coupled to a stochastic gravitational wave. (b) The phase-plane representation of the analogous single-parameter problem at a particular frequency Ω_0. Given vacuum input $|0\rangle$ with covariance matrix $\mathrm{diag}\left(\frac{1}{2}, \frac{1}{2}\right)$, the final quantum state $\hat{\rho}(\sigma)$ is a squeezed thermal state with covariance matrix $\mathrm{diag}\left(\frac{1}{2}, \frac{1}{2} + \sigma^2\right)$.

satisfying $[\hat{a}(t), \hat{a}^\dagger(t')] = \delta(t-t')$. The canonical quadratures of this mode are $\hat{x}(t) := \frac{1}{\sqrt{2}}[\hat{a}(t) + \hat{a}^\dagger(t)]$ and $\hat{p}(t) := \frac{1}{\sqrt{2}}[-i\hat{a}(t) + i\hat{a}^\dagger(t)]$, which obey $[\hat{x}(t), \hat{p}(t')] = i\delta(t-t')$. The generalized quadrature at angle θ is then defined as $\hat{x}_\theta(t) := \cos(\theta)\hat{x}(t) + \sin(\theta)\hat{p}(t)$ such that $\hat{x}_0(t) = \hat{x}(t)$ and $\hat{x}_{\pi/2}(t) = \hat{p}(t)$. Since the device is linear, these quadratures must obey

$$\hat{x}_\theta(t) = \hat{x}_\theta^{(0)}(t) + \int_{-\infty}^{\infty} \mathrm{d}t' \; \chi_\theta(t-t') y(t'), \tag{1}$$

where $\hat{x}_\theta^{(0)}(t)$ is the free solution (i.e., when no signal is present) and χ_θ is the linear susceptibility of the device [22,23]. In the frequency domain, all of the positive frequencies are independent since the device is linear, such that Eq. (1) becomes

$$\hat{x}_\theta(\Omega) = \hat{x}_\theta^{(0)}(\Omega) + \chi_\theta(\Omega) y(\Omega), \tag{2}$$

where we define the Fourier transform as $y(\Omega) := \int_{-\infty}^{\infty} \mathrm{d}t \; e^{i\Omega t} y(t)$. The frequency-domain quadratures are non-Hermitian but obey a conjugate symmetry $\hat{x}_\theta^\dagger(\Omega) = \hat{x}_\theta(-\Omega)$ and have the commutator $[\hat{x}(\Omega), \hat{p}^\dagger(\Omega')] = i2\pi\delta(\Omega - \Omega')$. Suppose that we measure $\hat{x}_\theta(t)$; then Eq. (2) implies that the observed output-referred power spectral density is

$$S_{x_\theta x_\theta}(\Omega) = S^{(0)}_{x_\theta x_\theta}(\Omega) + |\chi_\theta(\Omega)|^2 \, S_{yy}(\Omega), \tag{3}$$

where the total output-referred free noise $S^{(0)}_{x_\theta x_\theta}(\Omega)$ includes the quantum noise of the device, which we assume follows a stationary random process. In Eq. (3), the (one-sided) power spectral density $S_{x_\theta x_\theta}(\Omega)$ of the quadratures, which are stationary random processes in time, is defined as follows [24]

$$2\pi\delta(\Omega - \Omega') S_{x_\theta x_\theta}(\Omega) := \langle \{\hat{x}_\theta(\Omega), \hat{x}_\theta^\dagger(\Omega')\} \rangle, \tag{4}$$

where $\{\cdot,\cdot\}$ is the anticommutator and we assume that the quadratures have zero mean. For example, the output-referred power spectral density of the quantum shot noise of the vacuum state is $S^{(0)}_{x_\theta x_\theta}(\Omega) = 1$. Meanwhile, in Eq. (3), the (one-sided) power spectral density of the signal at each frequency is defined as [25]

$$S_{yy}(\Omega) = \lim_{T\to\infty} \frac{2}{T} \mathrm{E}\left[|y_T(\Omega)|^2\right], \tag{5}$$

where $y_T(\Omega) = \int_{-\infty}^{\infty} \mathrm{d}t \; e^{i\Omega t} y_T(t) = \int_0^T \mathrm{d}t \; e^{i\Omega t} y(t)$ is the Fourier transform of the signal $y_T(t) = [\Theta(t) - \Theta(t-T)]y(t)$ windowed to the finite time interval $(0, T)$ for some integration time T and Θ is the Heaviside function. Here, we take the expectation value over different realizations of the continuous random process describing the signal $y(t)$, which we assume has zero mean. Whenever we compute an expectation value with respect to the quantum state, e.g., $\langle\cdot\rangle$ in Eq. (4), we implicitly marginalize over the signal as well. We reserve the notation $\mathrm{E}[\cdot]$ for when we are only taking an expectation value with respect to the signal.

Let us mention a couple of well-known properties of the power spectral density $S_{yy}(\Omega)$. First, the average power P of the signal $y(t)$ over time, i.e., the variance $\mathrm{Var}[y]$ of the continuous random process, is given by

$$P = \frac{1}{2\pi} \int_0^\infty \mathrm{d}\Omega \; S_{yy}(\Omega). \tag{6}$$

And, second, provided that the signal $y(t)$ is ergodic, the power spectral density $S_{yy}(\Omega)$ equals the Fourier transform of the autocorrelation function $C_{yy}(t) = \lim_{T\to\infty}(1/T) \int_{-\infty}^{\infty} \mathrm{d}t' \; y_T(t'-t) y_T(t')$ by the Wiener-Khinchin theorem. These useful properties motivate why the power spectral density is an important quantity to estimate in an experiment.

Returning to the problem at hand, by using Eqs. (4)–(5), we can derive Eq. (3) from Eq. (2) [26]. We have thus far made only the following assumptions about the system: linearity, stationarity, and zero mean. We now make a further assumption about the tuning of the measurement device.

Some of our most sensitive devices of interest are optical interferometers for which, if the resonant optical cavities are all tuned and the Michelson interferometer is held at total destructive interference at the differential port, then a stochastic signal will only appear in one quadrature of the outgoing optical mode [27]. As such, we now restrict our attention to estimating the excess noise in one known quadrature of the outgoing mode. We will later generalize to sensing isotropic excess noise in both quadratures. Without further loss of generality, we assume that the stochastic signal only appears in the $\hat{p}(t)$ quadrature such that Eq. (2) implies that $\hat{x}(\Omega) = \hat{x}^{(0)}(\Omega)$ since $\chi_0 = 0$ and

$$\hat{p}(\Omega) = \hat{p}^{(0)}(\Omega) + \chi_{\frac{\pi}{2}}(\Omega) y(\Omega). \tag{7}$$

The output-referred power spectral density in Eq. (3) observed from measuring $\hat{p}(t)$ is then

$$S_{pp}(\Omega) = S^{(0)}_{pp}(\Omega) + \left|\chi_{\frac{\pi}{2}}(\Omega)\right|^2 S_{yy}(\Omega). \tag{8}$$

We consider running the experiment for a fixed finite integration time T that is long enough such that Eq. (5) implies that $S_{yy}(\Omega) \approx (2/T)\mathrm{E}[|y_T(\Omega)|^2]$ and Eq. (8) becomes

$$S_{pp}(\Omega) \approx S^{(0)}_{pp}(\Omega) + \frac{2}{T}\mathrm{E}\left[\left|\chi_{\frac{\pi}{2}}(\Omega) y_T(\Omega)\right|^2\right]. \tag{9}$$

We want to estimate the signal power $S_{yy}(\Omega)$ at each frequency Ω from measurements of the outgoing mode. We

have up to now considered continuous quadrature measurements. To understand other possible measurements, it is useful to first simplify this continuous-estimation problem.

B. Review of the deterministic case

Let us first review the deterministic case to establish an analogy to a toy model: continuum linear waveform estimation resembles estimating the displacement of a harmonic oscillator at each frequency. In the deterministic case, the signal $y(t)$ is a real-valued function and we want to estimate its Fourier component $y(\Omega)$, a complex number, at each frequency Ω. It is well established that this continuous-estimation problem reduces to a continuum of independent estimation problems, one for each frequency, and the fundamental quantum limits of this task are well understood [15–18]. We will discuss these sensitivity limits later, in Sec. V B. At a given frequency Ω, the canonical deterministic estimation problem in Eq. (7) is to estimate the displacement of $\hat{p}(\Omega)$ by $\chi_{\pi/2}(\Omega)y(\Omega)$ in the presence of noise from $\hat{p}^{(0)}(\Omega)$. Depending on the parametrization of $y(t)$, there are two independent real parameters of $y(\Omega)$ to estimate in general, e.g., its real and imaginary parts, or its amplitude and complex phase. The displacement-estimation problems at different frequencies are independent due to the linearity of the device: the covariance matrix of the quadratures at different frequencies is diagonal. Therefore, continuum-waveform estimation corresponds to estimating the displacements of a continuum of independent harmonic oscillators, each of which can be dealt with separately. The only assumptions that we have made to reach this conclusion are that the device is linear, stationary, zero-mean, and tuned.

C. Canonical noise-estimation problem

In the stochastic case, the signal $y(t)$ is now a continuous real-valued random process with variance given by Eq. (6). At each frequency Ω, we want to estimate $S_{yy}(\Omega)$, the contribution to the average power from that frequency component. Similarly to the deterministic case above, the continuum problem of estimating $S_{yy}(\Omega)$ at each frequency Ω is equivalent to a continuum of independent single-parameter harmonic oscillator estimation problems. In particular, the harmonic oscillators at different frequencies are independent due to linearity and at each frequency the signal appears as a displacement of the harmonic oscillator. Here, however, the displacement of $\hat{p}(\Omega)$ by $\chi_{\pi/2}(\Omega)y(\Omega)$ in Eq. (7) is stochastic, since $y(\Omega)$ is now a complex random variable. In particular, although the absolute value of $y(\Omega)$ is constrained by Eq. (5) in the limit of large T, its complex phase is uniformly distributed on $(0, 2\pi)$ since the signal $y(t)$ is zero-mean and stationary. Thus, we need to determine how the power spectral density $S_{yy}(\Omega)$ relates to the variance of the probability distribution describing these random displacements.

Let us first address a technicality about the frequency-domain quadratures. At a given frequency Ω, the quadratures $\hat{x}(\Omega)$ and $\hat{p}(\Omega)$ are non-Hermitian and comprise a total of four independent real degrees of freedom, whereas a harmonic oscillator has only two independent real degrees of freedom. This can be seen more clearly by decomposing the frequency-domain quadratures into their real and imaginary parts, which we label as $\hat{q}_j(\Omega)$ for $j = 1, 2, 3, 4$ and define as follows [28]:

$$\vec{\hat{q}}(\Omega) = \begin{bmatrix} \hat{q}_1(\Omega) \\ \hat{q}_2(\Omega) \\ \hat{q}_3(\Omega) \\ \hat{q}_4(\Omega) \end{bmatrix} = \sqrt{\frac{2}{T}} \begin{bmatrix} \mathrm{Re}[\hat{x}(\Omega)] \\ \mathrm{Re}[\hat{p}(\Omega)] \\ \mathrm{Im}[\hat{x}(\Omega)] \\ \mathrm{Im}[\hat{p}(\Omega)] \end{bmatrix}. \tag{10}$$

For a finite integration time T, this is equivalent to decomposing the outgoing mode into its cosine and sine phases at frequency Ω, since

$$\vec{\hat{q}}(\Omega) = \sqrt{\frac{2}{T}} \int_0^T \mathrm{d}t \begin{bmatrix} \cos(\Omega t)\hat{x}(t) \\ \cos(\Omega t)\hat{p}(t) \\ \sin(\Omega t)\hat{x}(t) \\ \sin(\Omega t)\hat{p}(t) \end{bmatrix}.$$

This implies that the commutator between $\hat{q}_1(\Omega)$ and $\hat{q}_2(\Omega')$ at two different frequencies Ω and Ω' equals

$$[\hat{q}_1(\Omega), \hat{q}_2(\Omega')] = i\left(\frac{\sin([\Omega - \Omega']T)}{[\Omega - \Omega']T} + \frac{\sin([\Omega + \Omega']T)}{[\Omega + \Omega']T}\right).$$

The second term vanishes, along with all other terms of size $\mathcal{O}(1/[\Omega T])$ henceforth, by assuming that we integrate for a long enough time $T \gg 1/\Omega$ given the lowest frequency of interest Ω. For frequencies that are further apart than the finite frequency resolution of $1/T$, i.e., that satisfy $[\Omega - \Omega']T \gg 1$, the first term also vanishes. However, for unresolved frequencies satisfying $[\Omega - \Omega']T \ll 1$ this commutator is approximately i, since the modes $\cos(\Omega t)$ and $\cos(\Omega' t)$ are not orthogonal on the finite interval $(0, T)$. We thus assume a discrete frequency sampling henceforth that is coarse enough, i.e., satisfying $[\Omega - \Omega']T \gg 1$, such that the harmonic oscillators at different frequencies commute. The number of such frequency bins is proportional to T and we will soon show that the noise and signal within a given bin are independent of T. This means that the total error in estimating, e.g., the power spectral density of a white-noise process will scale as $1/\sqrt{T}$ in amplitude units as expected.

We now focus on solving the estimation problem at a fixed frequency Ω. The four components of $\vec{\hat{q}}(\Omega)$ are Hermitian observables that resemble the quadratures of two harmonic oscillators, since $[\hat{q}_1(\Omega), \hat{q}_2(\Omega)] =$

$[\hat{q}_3(\Omega), \hat{q}_4(\Omega)] = i$, and all other commutators are zero. Since we assume that the signal and noise are zero-mean and stationary, the 4×4 covariance matrix, $\frac{1}{2}\langle\{\hat{q}_j(\Omega), \hat{q}_k(\Omega)\}\rangle$, is block diagonal such that the two harmonic oscillators are independent. For example, the covariance matrix for the vacuum state is $\frac{1}{2}\delta_{jk}$. By Eq. (7), these real and imaginary parts depend on the signal as follows:

$$\vec{\hat{q}}(\Omega) = \vec{\hat{q}}^{\,(0)}(\Omega) + \sqrt{\frac{2}{T}} \begin{bmatrix} 0 \\ \mathrm{Re}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right] \\ 0 \\ \mathrm{Im}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right] \end{bmatrix}. \tag{11}$$

The variances of the real and imaginary parts containing the zero-mean signal are then [29]

$$\begin{aligned} \mathrm{Var}\left[\hat{q}_2(\Omega)\right] &= \mathrm{Var}[\hat{q}_2^{(0)}(\Omega)] + \frac{2}{T}\mathrm{E}\left[\mathrm{Re}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right]^2\right], \\ \mathrm{Var}\left[\hat{q}_4(\Omega)\right] &= \mathrm{Var}[\hat{q}_4^{(0)}(\Omega)] + \frac{2}{T}\mathrm{E}\left[\mathrm{Im}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right]^2\right]. \end{aligned} \tag{12}$$

Let us simplify the excess variance due to the signal, i.e., the second term, starting with the real part as follows:

$$\begin{aligned} &\mathrm{E}\left[\mathrm{Re}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right]^2\right] \\ &\quad = \frac{1}{4}\mathrm{E}\left[\left(\chi_{\frac{\pi}{2}}(\Omega) y(\Omega) + \chi^*_{\frac{\pi}{2}}(\Omega) y^*(\Omega)\right)^2\right] \\ &\quad = \frac{1}{2}\mathrm{E}\left[\left|\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right|^2\right], \end{aligned}$$

where we have used the fact that $\mathrm{E}\left[y(\Omega)^2\right] = 0$, since the phase of $y(\Omega)$ is isotropically random. This excess variance in the real part of $\hat{p}(\Omega)$ equals half the excess power spectral density in Eq. (9) and the imaginary part is similar. The excess power spectral density is also equal to $|\chi_{\pi/2}(\Omega)|^2 S_{yy}(\Omega)$ by Eq. (8). Therefore, Eq. (12) may be rewritten as follows:

$$\begin{aligned} \mathrm{Var}\left[\hat{q}_2(\Omega)\right] &= \mathrm{Var}[\hat{q}_2^{(0)}(\Omega)] + \sigma^2, \\ \mathrm{Var}\left[\hat{q}_4(\Omega)\right] &= \mathrm{Var}[\hat{q}_4^{(0)}(\Omega)] + \sigma^2, \end{aligned} \tag{13}$$

where we have defined the common excess variance σ^2 and the gain $G_{py}(\Omega)$ as

$$\sigma^2 := \frac{1}{2} G_{py}(\Omega) S_{yy}(\Omega), \quad G_{py}(\Omega) := |\chi_{\frac{\pi}{2}}(\Omega)|^2. \tag{14}$$

The gain $G_{py}(\Omega)$ is also sometimes called the signal transfer function from y to $\hat{p}$, since it linearly relates the signal $S_{yy}(\Omega)$ to the observed spectral density $S_{pp}(\Omega)$ in Eq. (8) [30]. Since the two harmonic oscillators at each frequency are independent and the noise in each is equal due to stationarity, we can consider them separately to study the fundamental quantum limits on estimating $S_{yy}(\Omega)$ or, equivalently, estimating σ^2. Thus, we can restrict our attention to just the harmonic oscillator corresponding, e.g., to the real part with quadratures $\hat{q}_1(\Omega)$ and $\hat{q}_2(\Omega)$, which we relabel as $\hat{x}$ and $\hat{p}$ below for brevity.

For any initial state, we have shown that estimating the continuum $S_{yy}(\Omega)$ is equivalent to a continuum of independent copies (two for each frequency Ω) of the following canonical single-parameter estimation problem. Given the initial state $\hat{\rho}$ of a harmonic oscillator with canonical quadratures $\vec{\hat{x}} = (\hat{x}, \hat{p})^{\mathrm{T}}$ such that $[\hat{x}, \hat{p}] = i$, we wish to estimate σ^2 from measurements of the final state $\Lambda_\sigma(\hat{\rho})$ after the nonunitary quantum channel Λ_σ that encodes σ. (A quantum channel is a completely positive trace-preserving linear map between density matrices.) The signal-encoding channel Λ_σ represents a random displacement along the $\hat{p}$ quadrature and has the following Kraus representation:

$$\Lambda_\sigma(\hat{\rho}) = \int_{-\infty}^{\infty} \mathrm{d}\alpha\, p(\alpha) \hat{U}_\alpha \hat{\rho} \hat{U}_\alpha^\dagger, \tag{15}$$

where $p(\alpha) \sim \mathcal{N}(0, \sigma^2)$ is the weighting of the different displacement unitaries given by $\hat{U}_\alpha = \exp(i\alpha\hat{x})$, which displaces $\hat{p}$ to $\hat{p} + \alpha$. Here, we have assumed that the zero-mean signal follows a Gaussian random process such that the second moments in Eq. (13) determine the signal. Understanding this toy model of a single harmonic oscillator undergoing random displacements is the main focus of our work. We will solve this toy model in Secs. III–VI, before returning to the original stochastic waveform-estimation problem in Sec. VII.

We now provide two alternative expressions for the encoding channel Λ_σ that will prove useful later. First, in the basis of $\hat{x}$, Λ_σ is the following decoherence channel [31,32]:

$$\langle x|\Lambda_\sigma(\hat{\rho})|x'\rangle = e^{-\frac{1}{2}\sigma^2(x-x')^2} \langle x|\hat{\rho}|x'\rangle. \tag{16}$$

And, second, when $\sigma\hat{x}$ is small, we may expand Eq. (15) to obtain the infinitesimal channel

$$\Lambda_\sigma(\hat{\rho}) \approx \hat{\rho} + \sigma^2\left(\hat{x}\hat{\rho}\hat{x} - \frac{1}{2}\{\hat{x}^2, \hat{\rho}\}\right) + \mathcal{O}(\sigma^4\hat{x}^4), \tag{17}$$

where the odd terms vanish because $p(\alpha)$ is an even function in α.

Let us consider the Gaussian-state case. If the initial state $\hat{\rho}$ is a single-mode Gaussian with a 2×2 covariance matrix $\Sigma_{jk} = \frac{1}{2}\langle\{\vec{\hat{x}}_j^{(0)}, \vec{\hat{x}}_k^{(0)}\}\rangle$ and 2×1 mean vector

$\vec{\mu} = \langle \vec{\hat{x}} \rangle$, then the final state is also Gaussian with the covariance matrix $\Sigma + \mathrm{diag}(0, \sigma^2)$ and the same mean vector $\vec{\mu}$. For example, the vacuum state $|0\rangle\langle 0|$ with covariance matrix $\Sigma_0 = \mathrm{diag}(\frac{1}{2}, \frac{1}{2})$ becomes the squeezed thermal state $\hat{\rho}(\sigma) = \Lambda_\sigma(|0\rangle\langle 0|)$ with covariance matrix $\mathrm{diag}(\frac{1}{2}, \frac{1}{2} + \sigma^2)$, as shown in Fig. 1(b). We thus call Λ_σ a Gaussian-encoding channel, since it sends Gaussian states to Gaussian states.

Our goal now is to solve this canonical noise-estimation problem by understanding the optimal protocol for estimating σ^2. As we will see, despite Λ_σ being a Gaussian-encoding channel, the optimal protocol will turn out to be to prepare non-Gaussian initial states and perform non-Gaussian measurements.

D. Review of Fisher information

We now review our main tools for addressing this single-parameter estimation problem: the concepts of classical and quantum Fisher information (FI) [33–35].

Suppose that a real parameter of interest θ is encoded in a quantum state $\hat{\rho}(\theta)$ and that we estimate θ by performing a given measurement [i.e., a positive operator-valued measure (POVM)] with associated probability distribution $p(x|\theta)$. The minimal MSE $\Delta^2\theta$ of unbiased estimation of θ from M independent and identically distributed measurement results satisfies the *classical Cramér-Rao bound (CCRB)*, $\Delta^2\theta \geq (1/M)[\mathcal{I}_C(\theta)]^{-1}$, where $\mathcal{I}_C(\theta)$ is the *classical Fisher information (CFI)*, given by

$$\mathcal{I}_C(\sigma) = \int_{-\infty}^{\infty} \mathrm{d}x \, \frac{[\partial_\theta p(x|\theta)]^2}{p(x|\theta)}. \tag{18}$$

This bound can be attained by maximum likelihood estimation in the asymptotic limit such that the central limit theorem applies. In general, saturating the CFI may require a parameter that is sufficiently small or has a narrow enough prior, as well as the ability to perform a large number of independent measurements. A useful property of the CFI is that if $p(x|\theta)$ and $p(y|\theta)$ are independent distributions, e.g., describing two separate measurements, then the total CFI from observing one outcome from each distribution is simply the sum of the individual CFIs.

This may not be the optimal measurement, however, for extracting the maximal information about θ from $\hat{\rho}(\theta)$. For single-parameter estimation, the *quantum Fisher information (QFI)* is the CFI maximized over all possible measurements (POVMs), $\mathcal{I}_Q(\theta) = \sup \mathcal{I}_C(\theta)$. In terms of the eigenvalues p_j and eigenvectors $|\phi_j\rangle$ of the final state $\hat{\rho}(\theta) = \sum_j p_j |\phi_j\rangle\langle\phi_j|$, the QFI can be shown to be

$$\mathcal{I}_Q(\theta) = \sum_{j,k} \frac{2}{p_j + p_k} \left| \langle \phi_j | \partial_\theta \hat{\rho}(\theta) | \phi_k \rangle \right|^2$$

where the sum runs over only j and k such that $p_j + p_k > 0$. For example, if a parameter θ is encoded by a unitary $\exp(-i\theta\hat{H})$ applied to a pure state $|\psi\rangle$, then the QFI is $\mathcal{I}_Q(\theta) = 4\mathrm{Var}_{|\psi\rangle}[\hat{H}]$ independent of θ.

The *quantum Cramér-Rao bound (QCRB)*, $\Delta^2\theta (\geq 1/M)[\mathcal{I}_Q(\theta)]^{-1}$, provides the fundamental minimal MSE that can be achieved by maximum likelihood estimation from the M outcomes of the optimal measurement scheme. Similarly, the minimal fractional MSE with respect to a parameter θ is bounded too, $\Delta^2\theta/\theta^2 \geq (1/M)[\theta^2\mathcal{I}_Q(\theta)]^{-1}$. The analog of the additivity of the CFI for independent distributions is that the QFI for a product state $\hat{\rho}_1(\theta) \otimes \hat{\rho}_2(\theta)$ is simply the sum of the individual QFIs.

E. Estimating standard deviation versus variance

For our single-parameter problem, we choose to estimate the standard deviation σ. This is equivalent to estimating the variance σ^2, since $\sigma \geq 0$. The QFIs with respect to σ and σ^2 are related by $\mathcal{I}_Q(\sigma) = 4\sigma^2\mathcal{I}_Q(\sigma^2)$ through the chain rule. Although $\mathcal{I}_Q(\sigma^2)$ may diverge and $\mathcal{I}_Q(\sigma)$ stay finite as $\sigma \to 0$, the limiting behavior of the fractional MSEs is consistent since, in that event, $[\sigma^4\mathcal{I}_Q(\sigma^2)]^{-1} \propto [\sigma^2\mathcal{I}_Q(\sigma)]^{-1} \to \infty$.

III. LOSSLESS VACUUM LIMIT

The task of estimating σ from Λ_σ in Eq. (15), in the absence of loss when the input state is the vacuum, has previously been studied in Refs. [19,20,36]. We briefly discuss this ideal lossless regime here for comparison.

Suppose that we apply Λ_σ to $\hat{\rho} = |0\rangle\langle 0|$ such that the final covariance matrix is $\Sigma = \mathrm{diag}(\frac{1}{2}, \frac{1}{2} + \sigma^2)$. The QFI for a signal encoded solely in the covariance matrix of a single-mode Gaussian state is [37]

$$\mathcal{I}_Q(\sigma) = \frac{\mathrm{Tr}\left[(\Sigma^{-1}\partial_\sigma\Sigma)^2\right]}{2(1+\gamma^2)} + \frac{2(\partial_\sigma\gamma)^2}{1-\gamma^4}, \tag{19}$$

where $\gamma = \det(2\Sigma)^{-\frac{1}{2}}$ is the purity of the Gaussian state. For the vacuum state, Eq. (19) implies that the QFI is

$$\mathcal{I}_Q(\sigma) = \frac{2}{1+\sigma^2} \xrightarrow[\sigma\to 0]{} 2. \tag{20}$$

We will write the important limit of $\lim_{\sigma\to 0}\mathcal{I}_Q(\sigma)$ as $\mathcal{I}_Q(\sigma = 0)$ henceforth. Note that it is nonzero for the QFI.

We want to know what measurement will saturate the QFI in Eq. (20). A homodyne measurement of $\hat{p}$ is a Gaussian measurement for which the CFI with respect to the total standard deviation $\varsigma = \sqrt{\frac{1}{2} + \sigma^2}$ is $\mathcal{I}_C^{\hat{p}}(\varsigma) = 2/\varsigma^2$. This implies that $\mathcal{I}_C^{\hat{p}}(\sigma) = 2\sigma^2/\left(\frac{1}{2} + \sigma^2\right)^2$ such that the CFI for quadrature measurement saturates the QFI in Eq. (20) in the classical regime of $\sigma \gg \frac{1}{\sqrt{2}}$ but vanishes in the quantum regime of $\sigma \to 0$, where $\mathcal{I}_C^{\hat{p}}(\sigma) \approx 8\sigma^2$,

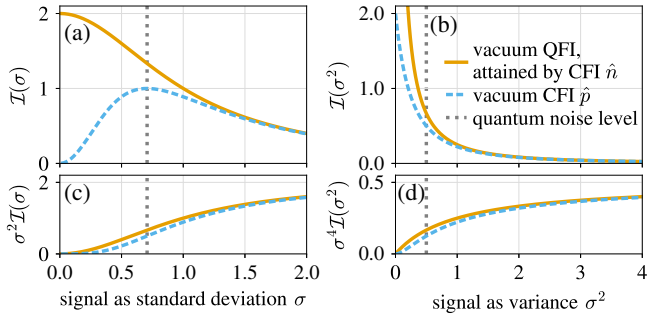

FIG. 2. The Fisher information (FI) with respect to (a) standard deviation σ or (b) variance σ^2 for the single-parameter estimation problem in the lossless vacuum limit. The level of quantum noise for the initial vacuum state, shown by the dotted gray vertical line, is equal to $\mathrm{Var}[\hat{p}] = \frac{1}{2}$. A quadrature measurement $\hat{p}$, with CFI shown in the dashed blue curve, is optimal for signals well above this level. A number-resolving measurement $\hat{n}$, however, is optimal—its CFI attains the QFI shown in the solid orange curve—for all σ including as $\sigma \to 0$. In this limit, the fractional MSE with respect to (c) σ or (d) σ^2 diverges regardless of the measurement scheme, since $\theta^2\mathcal{I}_Q(\theta)$ converges to zero.

as shown in Fig. 2(a). If we estimate σ^2 instead of σ, as shown in Fig. 2(b), then the behavior of the fractional MSE is the same as discussed in Sec. II E and shown in Fig. 2(c). The conventional quadrature measurement, therefore, is highly inefficient in the relevant limit of $\sigma \to 0$. Intuitively, the small change in the variance induced by the weak signal $\sigma^2 \ll \frac{1}{2}$ is masked by the vacuum quantum noise, leading to this vanishing sensitivity.

In comparison, a number-resolving measurement of $\hat{n} = \hat{a}^\dagger\hat{a}$ can be shown to saturate the QFI for all σ. Since $\hat{U}_\alpha|0\rangle = |\alpha'\rangle$ is a coherent state with amplitude $\alpha' = i\alpha/\sqrt{2}$, we can compute the probabilities in the number basis after the encoding in Eq. (15) as

$$p(n) = \int_{-\infty}^{\infty} \mathrm{d}\alpha \ \left|\langle n|\alpha'\rangle\right|^2 p(\alpha) = \frac{(2n)!}{2^{2n}(n!)^2}\frac{\sigma^{2n}}{\left(\sigma^2+1\right)^{n+\frac{1}{2}}}. \tag{21}$$

Then, the CFI for number measurement, calculated as $\mathcal{I}_C^{\hat{n}}(\sigma) = \sum_{n=0}^{\infty}[\partial_\sigma p(n)]^2/p(n)$ by Eq. (18), equals the QFI in Eq. (20), as shown in Fig. 2. Intuitively, if there is no signal, then the vacuum state remains an eigenstate of $\hat{n}$ and no particles are counted. If a particle is counted, then it must come from the signal. This is unlike the quadrature measurement, where some variance in the measurement results remains due to the quantum noise even if there is no signal.

Let us illustrate how much worse quadrature measurement is compared to the optimal number-resolving measurement. If the signal is comparable to the quantum noise, i.e., $\sigma = \frac{1}{\sqrt{2}}$, then the QFI is $\frac{4}{3}$ compared to the quadrature measurement CFI of 1. Instead, if the signal is 10 times (100 times) smaller than the quantum noise in amplitude units, i.e., $\sigma = \frac{1}{10\sqrt{2}}$ ($\sigma = \frac{1}{100\sqrt{2}}$), then the QFI is roughly 51 times (5001 times) greater than the CFI. Correspondingly, suppose that we want to achieve an MSE $\Delta^2\sigma$ below some target value. If $\sigma = \frac{1}{10\sqrt{2}}$, e.g., then it takes only 2% of the number of quadrature measurements required to attain the target value if we instead use number-resolving measurements (recall that the Cramér-Rao Bound on the MSE is $(1/M)[\mathcal{I}(\sigma)]^{-1}$ given M measurements).

This estimation problem is analogous to quantum superresolution in optical imaging [20,38,39] and spectroscopy [40,41]. These resolution problems are characterized by having vanishing signal: $\partial_\theta\hat{\rho} \to 0$ as $\theta \to 0$, where θ is the parameter of interest. This vanishing signal leads to a "Rayleigh curse," where the CFI of the naive measurement, e.g., image-plane photon counting in the case of optical imaging, also vanishes as $\theta \to 0$. The QFI, however, remains positive in this limit and is attained by the CFI of a particular nonstandard measurement, e.g., spatial-mode demultiplexing in the case of optical imaging. Analogously, we define the *Rayleigh curse* to refer to any scenario in which the FI converges to zero in the limit of zero signal. We have seen above that here the Rayleigh curse arises for the CFI of the naive quadrature measurement but can be avoided by performing a number measurement.

IV. LOSS AND CLASSICAL NOISE

Realistically, the quantum state will experience noise channels before and after the encoding such that the noiseless channel Λ_σ becomes the noisy channel Λ'_σ. We restrict our attention to Gaussian noise channels. We want to understand how these imperfections limit our ability to estimate σ and whether the Rayleigh curse can still be avoided to increase QFI in the small-signal limit.

A. Gaussian states

We first consider preparing a Gaussian initial state in the limit of $\sigma \to 0$. For Gaussian noise channels, $\partial_\sigma\Lambda'_\sigma \to 0$ as $\sigma \to 0$ and information about σ is encoded only in the covariance matrix of the output Gaussian state. Whether the Rayleigh curse reappears is addressed by the following claim.

Claim 1. For any Gaussian state such that a parameter σ is encoded only on its covariance matrix $\Sigma(\sigma)$ and $\lim_{\sigma\to 0}\partial_\sigma\Sigma = 0$, then $\mathcal{I}_Q(\sigma = 0) \neq 0$ (i.e., the state does not exhibit the Rayleigh curse) if and only if there exists a symplectic eigenvalue of Σ equal to $\frac{1}{2} + k\sigma^2$ for some constant $k > 0$.

This is the Gaussian version of the quantum resolution criterion in Ref. [40]. The proof of this claim is given in Appendix A. For a single-mode Gaussian state, this claim implies that overcoming the Rayleigh curse is possible only if $\Lambda'_{\sigma=0}(\rho)$ is pure. In that case, the QFI comes only from the purity [the second term in Eq. (19)]: the QFI is nonvanishing in the limit $\sigma \to 0$, and equal to $\mathcal{I}_Q(\sigma = 0) = 2k'$, if and only if the purity is $\gamma = 1 - k'\sigma^2$ for some constant k'. For a multimode Gaussian state in the limit $\sigma \to 0$, at least one of the modes needs to be pure for the QFI not to vanish.

B. Loss channel

We apply Claim 1 to a couple of examples of common Gaussian noise channels, assuming an initial vacuum state. First, consider the "pure" or "cold" loss channel $\Lambda_\eta^{\text{loss}}$, which models, e.g., losses due to coupling with a zero-temperature bath. This can be modeled as a beam-splitter operation with an ancillary vacuum mode that is then traced out as shown in Fig. 3. For a loss $\eta \in (0, 1)$, this channel has the following Kraus representation [42]:

$$\Lambda_\eta^{\text{loss}}(\hat{\rho}) = \sum_{n=0}^{\infty} \hat{K}_n \hat{\rho} \hat{K}_n^\dagger,$$

where the Kraus operators are

$$\hat{K}_n = \frac{1}{\sqrt{n!}} \left(\sqrt{\eta}\hat{a}\right)^n (1-\eta)^{\frac{1}{2}(\hat{a}^\dagger \hat{a} - n)}.$$

Note that here η is the loss and $1 - \eta$ is the efficiency. The effect of this channel on the mean vector and covariance matrix of a single-mode Gaussian state is as follows: $\vec{\mu} \mapsto \sqrt{1-\eta}\vec{\mu}$ and $\Sigma \mapsto (1-\eta)\Sigma + \eta\Sigma_0$, where $\Sigma_0 = \text{diag}\left(\frac{1}{2}, \frac{1}{2}\right)$. If the initial state is a vacuum state, then the final state in the limit of $\sigma \to 0$ is pure and there is no Rayleigh curse. The precise nonzero value of the vacuum QFI depends on whether the loss occurs before or after the encoding. If the loss occurs before the encoding, then the QFI is unchanged from $2/(1+\sigma^2)$, since the vacuum $|0\rangle$ is stable under the loss channel $\Lambda_\eta^{\text{loss}}$. But if the loss occurs after the encoding, then the QFI degrades to $2(1-\eta)/(1+(1-\eta)\sigma^2)$, since the state after the encoding is no longer vacuum and therefore is not stable under the loss channel.

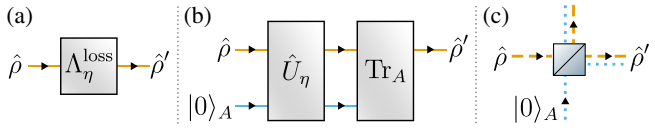

FIG. 3. The quantum state $\hat{\rho}'$ after a loss can be thought of as: (a) the result, $\hat{\rho}' = \Lambda_\eta^{\text{loss}}(\hat{\rho})$, of a nonunitary quantum channel $\Lambda_\eta^{\text{loss}}$; (b) the result of a beam-splitter unitary $\hat{U}_\eta$ with an ancillary vacuum mode $|0\rangle_A$ that is then traced out; or (c) for an optical system, the state at one of the output ports of a fictitious beam splitter with a vacuum mode.

In this work, we will consider different initial states that will not be stable under the loss channel, unlike the vacuum. We will focus on the impact of a known loss η occurring before the encoding, which now will affect the QFI, and assume that no loss occurs after the encoding. We address the effect of noise channels occurring after the encoding in Appendix B 1, where we show that measurement noise can, in theory, be overcome by using a suitable control unitary. We analyze the implications of an unknown loss in Appendix B 2, where we justify neglecting this effect given a vacuum input state.

C. Classical noise channel

We also consider a Gaussian noise channel $\Lambda_{\Sigma_C}^{\text{noise}}$ that models a source of uncorrelated classical noise, e.g., thermal fluctuation processes in optics or a nonzero temperature of microwave resonators. For M modes, this channel is a random-displacement channel

$$\Lambda_{\Sigma_C}^{\text{noise}}(\hat{\rho}) = \int_{\mathbb{C}^M} \mathrm{d}\vec{\alpha}\, p(\vec{\alpha}) \hat{D}(\vec{\alpha}) \hat{\rho} \hat{D}(\vec{\alpha})^\dagger, \tag{22}$$

where $p(\alpha) \sim \mathcal{N}(\mathbf{0}, \frac{1}{2}\Sigma_C)$ for a positive-semidefinite $2M \times 2M$ matrix Σ_C and the M-mode displacement operator is $\hat{D}(\alpha) = \prod_{i=1}^M \exp(\alpha_i \hat{a}_i^\dagger - \alpha_i^* \hat{a}_i)$, where $\hat{a}_i$ is the annihilation operator for the ith mode. This classical noise channel acts on Gaussian states as an additive noise source: $\Sigma \mapsto \Sigma + \Sigma_C$. [Note that $\Lambda_\sigma = \Lambda_{\text{diag}(0,\sigma^2)}^{\text{noise}}$ in Eq. (15).]

For a given single-mode Gaussian state with fixed $\langle \hat{n} \rangle = N$, a nonzero Σ_C will make the final state mixed and always lead to a Rayleigh curse, as $\sigma \to 0$ by Claim 1. For example, if $\Sigma_C = \text{diag}\left(\sigma_x^2, \sigma_p^2\right)$ with vacuum input, then

$$\mathcal{I}_Q(\sigma) = \frac{2\sigma^2 \left(1 + 2\sigma_x^2\right)^2}{\kappa\,(\kappa + 1)}, \tag{23}$$

where

$$\kappa = \sigma^2 + \sigma_x^2 + \sigma_p^2 + 2\sigma_x^2 \left(\sigma^2 + \sigma_p^2\right). \tag{24}$$

As long as one of σ_x and σ_p is nonzero, then $\mathcal{I}_Q(\sigma = 0) = 0$ here. In particular, for the isotropic case $\sigma_C := \sigma_x = \sigma_p$, the Rayleigh curse arises when the signal is dominated by the classical noise, $\sigma \ll \sigma_C$.

Classical noise changes the optimal measurement for the vacuum input-state case, such that number measurement alone no longer saturates the QFI. Instead, the optimal measurement is to squeeze $\Sigma \mapsto \text{diag}\left(e^{-2r}, e^{2r}\right)\Sigma$ with $e^{2r} = 1 + 2\sigma_x^2$ immediately after

the encoding process. The squeezing level is chosen such that $\Sigma = \mathrm{diag}\left(\frac{1}{2}+\sigma_x^2, \frac{1}{2}+\sigma^2+\sigma_p^2\right)$ becomes $\mathrm{diag}\left(\frac{1}{2}, \frac{1}{2}+\kappa\right)$. After this additional squeezing operation, we then perform a number measurement. The gain in the CFI from this additional squeezing operation in the limit of small classical noise compared to direct number measurement is $1+4\sigma_x^2$; hence it is marginal given small σ_x but it is still highly favorable compared to quadrature measurement.

Let us discuss how to determine, e.g., σ_p for a given experiment. The total output-referred free noise $S_{pp}^{(0)}(\Omega)$ in Eq. (3) is the sum of the output-referred quantum noise $S_{pp}^{Q}(\Omega)$ and the output-referred classical noise $S_{pp}^{C}(\Omega)$ such that

$$S_{pp}(\Omega) = S_{pp}^{Q}(\Omega) + S_{pp}^{C}(\Omega) + G_{py}(\Omega)S_{yy}(\Omega). \tag{25}$$

For example, the output-referred quantum noise is $S_{pp}^{Q}(\Omega) = S_{pp}^{Q,\mathrm{vac}}(\Omega) = 1$ if the quantum noise spectrum is dominated by the quantum shot noise from the vacuum, which we denote as $S_{pp}^{Q,\mathrm{vac}}(\Omega)$ (i.e., if the input state is vacuum and the device has no effects, such as quantum radiation pressure noise, that modify the quantum state if there is zero signal). Similarly to how σ is given by Eq. (14), the parallel classical noise is thus $\sigma_p^2 := \frac{1}{2}S_{pp}^{C}(\Omega)$ and, analogously, the perpendicular classical noise is $\sigma_x^2 := \frac{1}{2}S_{xx}^{C}(\Omega)$. (More generally, Σ_C equals $1/2$ times the classical noise contributions to the matrix of cross spectra of $\hat{x}(\Omega)$ and $\hat{p}(\Omega)$.) Furthermore, let us relate, e.g., σ_p to a common set of performance metrics for interferometers: the signal-referred noise spectral densities. These metrics are obtained by dividing Eq. (25) by the gain $G_{py}(\Omega)$ and are given in power units as follows:

$$\begin{aligned} \mathcal{S}(\Omega) &:= \frac{S_{pp}(\Omega)}{G_{py}(\Omega)} = \mathcal{S}^{Q}(\Omega) + \mathcal{S}^{C}(\Omega) + S_{yy}(\Omega), \\ \mathcal{S}^{Q}(\Omega) &:= \frac{S_{pp}^{Q}(\Omega)}{G_{py}(\Omega)}, \quad \mathcal{S}^{C}(\Omega) := \frac{S_{pp}^{C}(\Omega)}{G_{py}(\Omega)}. \end{aligned} \tag{26}$$

Here, $\mathcal{S}^{Q}(\Omega)$ is the signal-referred quantum noise, $\mathcal{S}^{C}(\Omega)$ is the signal-referred classical noise, and $\mathcal{S}^{(0)}(\Omega) = \mathcal{S}^{Q}(\Omega) + \mathcal{S}^{C}(\Omega)$ is the total signal-referred free noise. For example, for gravitational-wave observatories, $\mathcal{S}^{(0)}(\Omega)$ is the strain sensitivity to deterministic gravitational waves. If the total free-noise spectrum is dominated by the quantum shot noise from the vacuum, which is the case for gravitational-wave observatories at kilohertz frequencies, then $\mathcal{S}^{(0)}(\Omega) = \mathcal{S}^{Q,\mathrm{vac}}(\Omega) = 1/G_{py}(\Omega)$, where $\mathcal{S}^{Q,\mathrm{vac}}(\Omega) = S_{pp}^{Q,\mathrm{vac}}(\Omega)/G_{py}(\Omega)$. This provides the following simple relationship to determine the gain $G_{py}(\Omega)$ from existing tuned interferometer designs that meet these assumptions:

$$G_{py}(\Omega) = \mathcal{S}^{Q,\mathrm{vac}}(\Omega)^{-1}. \tag{27}$$

The classical noise may then be determined from the signal-referred spectra, e.g., $\sigma_p^2 := \frac{1}{2}\mathcal{S}^{C}(\Omega)/\mathcal{S}^{Q,\mathrm{vac}}(\Omega)$.

In summary, classical noise limits the QFI for small signals when preparing the vacuum state. Unless stated otherwise, we assume henceforth that the classical noise is negligible, i.e., that the signal is dominant, $\sigma \gg \sigma_C$. This is motivated by certain applications, discussed later in Sec. VII B, for which the search for stochastic signals is limited by imperfections from decoherence—quantum backgrounds—rather than classical backgrounds. We will revisit nonzero classical noise and whether its impact can be avoided by preparing different initial states in Sec. V G.

V. OPTIMAL INITIAL STATE

We now consider the initial state and measurement scheme that comprise the optimal protocol for sensing the signal σ encoded by Λ_σ with and without losses. In particular, we want to know whether entangled resources and collective measurements are necessary.

A. Review of channel quantum Fisher information

Building on Sec. II D, here, we introduce our tools for determining the optimal initial state: the channel QFI and extended-channel QFI [43,44].

Given an initial state $|\psi\rangle$ and a channel Λ_θ that encodes a parameter θ in the final state $\hat{\rho}(\theta) = \Lambda_\theta(|\psi\rangle\langle\psi|)$, then let the QFI with respect to θ be denoted as $\mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta)$. The QFI from the optimal initial state $|\psi\rangle$ is called the *channel QFI (CQFI)* of Λ_θ and is given by

$$\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta) = \sup_{|\psi\rangle} \mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta), \tag{28}$$

where we emphasize that no ancilla is allowed when calculating the CQFI. From the convexity of the QFI, it suffices to optimize over pure initial states.

If we prepare an initial state $|\Psi\rangle$ that is also entangled with some ancilla, e.g., a TMSV state, then we might improve the estimation of θ. Let the joint channel be $\Lambda_\sigma \otimes \Lambda_A$, where Λ_A is some channel that acts on the ancilla. In the noiseless-ancilla case, $\Lambda_A = 1_A$ is the identity. In this ideal case, the CQFI of the joint channel $\Lambda_\sigma \otimes 1_A$ is called the *extended-channel QFI (ECQFI)* of Λ_σ and is given by

$$\mathcal{I}_Q^{\Lambda_\theta}(\theta) = \sup_{|\Psi\rangle} \mathcal{I}_Q^{(\Lambda_\theta\otimes 1_A)(|\Psi\rangle\langle\Psi|)}(\theta). \tag{29}$$

The ECQFI is the maximum amount of information about the parameter θ that can be extracted after the channel Λ_θ acts on the quantum device. Note that the inequality $\mathcal{I}_Q^{\Lambda_\theta}(\theta) \geq \mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta)$ between the ECQFI and the

CQFI always holds. If, as in some cases, $\mathcal{I}_Q^{\Lambda_\theta}(\theta)$ is strictly larger than $\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta)$, then entanglement with an ancilla is a required resource for optimal signal extraction [45].

Realistically, however, we expect that the ancilla should also experience some loss such that $\Lambda_A = \Lambda_{\eta_A}^{\text{loss}}$, where η_A is some ancilla loss, rather than $\Lambda_A = 1_A$. The CQFI of this joint channel $\Lambda_\sigma \otimes \Lambda_{\eta_A}^{\text{loss}}$ then represents the maximum information feasibly available. The key questions to ask are then whether the CQFI and ECQFI of $\Lambda_\sigma \otimes \Lambda_{\eta_A}^{\text{loss}}$ are equal and, similarly, whether the CQFI and ECQFI of Λ_σ are equal. If the latter is true, then all four of the CQFI and ECQFI of $\Lambda_\sigma \otimes \Lambda_{\eta_A}^{\text{loss}}$ and the CQFI and ECQFI of Λ_σ are equal and the ultimate sensitivity limit can be achieved without using an ancilla. This will turn out to be the case in our work (see Sec. V F below).

When exploring these limits, we can also impose other physical restrictions. For example, since the unconstrained ECQFI $\mathcal{I}_Q^{\Lambda_\theta}(\theta)$ might be unbounded at $\theta = 0$ in the lossless case, we can constrain the average energy $\langle \hat{H} \rangle \propto \langle \hat{n} \rangle + \frac{1}{2}$ of the initial state of the harmonic oscillator. Let the ECQFI in Eq. (29) constrained to initial states $|\Psi\rangle$ with $\langle \hat{n} \rangle = N$ average occupation number per mode be denoted $\mathcal{I}_Q^{\Lambda_\theta,N}(\theta)$. (Note that while $\langle \hat{n} \rangle \leq N$ may be a more natural constraint, the bounds that we find are always nondecreasing in N such that it suffices to consider $\langle \hat{n} \rangle = N$.)

We need to therefore calculate the ECQFI of our channel Λ_σ with respect to σ to determine the optimal initial state and whether entanglement is a required resource, with and without noise channels on the system and ancilla.

B. Review of the deterministic case

Before proceeding to our case of a random-displacement channel, we briefly review the case of a deterministic displacement channel to establish the similarities and differences between the two cases.

Consider a deterministic displacement of the state $\hat{\rho}$ by μ along $\hat{p}$ such that we want to estimate μ from measurements of $\hat{U}_\mu \hat{\rho} \hat{U}_\mu^\dagger$, where $\hat{U}_\mu = \exp(i\mu\hat{x})$. In the lossless case, if $\hat{\rho} = |\psi\rangle\langle\psi|$ is pure, then the QFI is $\mathcal{I}_Q(\mu) = 4\text{Var}_{|\psi\rangle}[\hat{x}]$. Since $\text{Var}_{|\psi\rangle}[\hat{x}]$ can be made arbitrarily large, the unconstrained ECQFI is unbounded. As such, we constrain the initial state $|\psi\rangle$ to have N average occupation number per mode, $\langle \hat{n} \rangle = N$. The maximum value of $\text{Var}_{|\psi\rangle}[\hat{x}]$ given the constraint of $\langle \hat{n} \rangle = N$ is attained by an SMSV state, in which case $\text{Var}_{|\psi\rangle}[\hat{x}]$ is equal to $\xi_N := N + \frac{1}{2} + \sqrt{N(N+1)}$. Note that $\xi_N \to 2N$ as $N \to \infty$. The ECQFI is then

$$\mathcal{I}_Q^{\hat{U}_\mu,N}(\mu) = 4\xi_N \xrightarrow[N\to\infty]{} 8N, \tag{30}$$

which is achieved by preparing an SMSV state and, e.g., measuring the quadrature $\hat{p}$. Another optimal measurement is to instead antisqueeze after the encoding and then perform a number measurement $\hat{n}$. We emphasize that this asymptotic scaling with N is the fundamental limit for lossless deterministic displacement sensing. The Heisenberg limit of N^2 for phase estimation cannot be achieved for displacements.

Realistically, however, the state will experience some losses such that the total channel is not unitary. Suppose that the state encounters a loss η before the encoding such that the total channel is $\Lambda'_\mu(\rho) = \hat{U}_\mu \Lambda_\eta^{\text{loss}}(\hat{\rho}) \hat{U}_\mu^\dagger$. In this lossy case, Ref. [18] has shown that an SMSV state is still the optimal initial state and attains the following ECQFI given the constraint of $\langle \hat{n} \rangle = N$:

$$\mathcal{I}_Q^{\Lambda'_\mu,N}(\mu) = \frac{4\xi_N}{\eta(2\xi_N - 1) + 1} \xrightarrow[N\to\infty]{} \frac{2}{\eta}. \tag{31}$$

The high-energy limit of the deterministic ECQFI is then bounded by $2/\eta$, which cannot be surpassed using any initial state. For example, this result is known in the gravitational-wave-observatory context where, to sense deterministic gravitational waves in the presence of optical losses, it is optimal to inject a squeezed optical state into the differential port of the Michelson interferometer and perform a quadrature measurement [46], as is presently done [47–50].

C. Lossless case

We now return to studying the ECQFI of the random-displacement channel Λ_σ with respect to σ. We prove the following claim about the optimal protocol.

Claim 2. In the lossless case, the ECQFI constrained to initial states with $\langle \hat{n} \rangle = N$ per mode is

$$\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma) = \frac{4}{2\sigma^2 + \xi_N^{-1}}, \tag{32}$$

which is saturated by preparing an initial SMSV state and performing a number-resolving measurement after first antisqueezing. Entanglement with an ancilla is not a required resource.

The proof of Claim 2 is given in Appendix C 2. We use an established technique of bounding the ECQFI using a sequence of purifications of Λ_σ by Uhlmann's theorem [18,43,44]. We find the following upper bound on the ECQFI:

$$\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma) \leq \frac{4}{2\sigma^2 + \text{Var}[\hat{x}]^{-1}}, \tag{33}$$

where $\text{Var}[\hat{x}]$ is calculated with respect to the initial state. Minimizing Eq. (33) over the initial state is equivalent to

optimizing $\mathrm{Var}[\hat{x}]$ given the constraint of $\langle\hat{n}\rangle = N$, which is attained by an SMSV state. By calculating the QFI using Eq. (19), we observe that an SMSV state is optimal for any σ and saturates the ECQFI in Eq. (32).

Claim 2 implies that the behavior of the ECQFI depends on the ratio between N and $1/4\sigma^2$. For $N \ll 1/4\sigma^2$, $\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma)$ grows linearly as $8N$ similarly to Eq. (30) but if $N \gg 1/4\sigma^2$, then $\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma)$ converges to $2/\sigma^2$.

We comment on the limit of vanishingly small signals $\sigma \to 0$ such that the upper bound in Eq. (33) becomes $4\mathrm{Var}[\hat{x}]$ and is tight for any initial state in this limit [51]. (Curiously, this equals the QFI for deterministic displacements in Sec. V B.) In Fig. 4(a), for this small-signal limit of $\sigma \to 0$, we compare the ECQFI attained by an SMSV state versus the average number $\langle\hat{n}\rangle$ to the QFI of other states, such as coherent, TMSV, Fock, Schrödinger's cat, and finite-energy Gottesman-Kitaev-Preskill (GKP) grid states [52]. (We review finite-energy GKP states in Appendix H.) In the high-energy limit of $\langle\hat{n}\rangle \to \infty$; cat states also attain the ECQFI of $8N$; TMSV, finite-energy GKP states, and Fock only grow as $4N$; and coherent states remain at the vacuum level of 2, since displacements commute with the encoding channel Λ_σ. In this limit of $\sigma \to 0$, an optimal measurement of the final state $\hat{\rho}$ is to project it onto the initial state $|\Psi\rangle\langle\Psi|$, which can be implemented by an "echo protocol" [53]. For example, it suffices to directly perform a number-resolving measurement of the SMSV in this vanishing-signal limit of $\sigma \to 0$. For arbitrary $\sigma > 0$, however, we need to first antisqueeze and then perform a number-resolving measurement as stated in Claim 2.

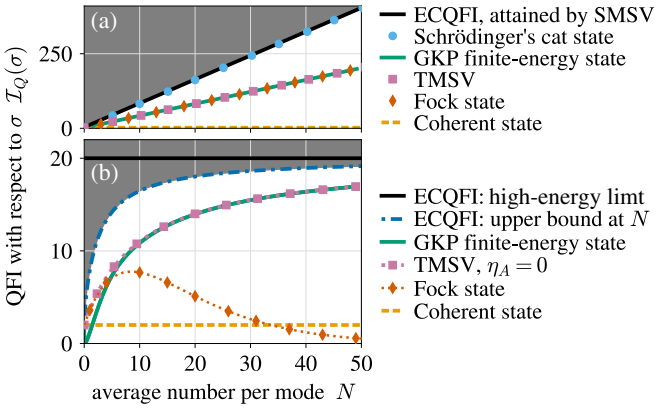

FIG. 4. The QFI from preparing different initial states, indicated in the legend, versus the initial average occupation number per mode in (a) the lossless case and (b) the case of a loss of $\eta = 0.1$ occurring before the encoding with $\sigma = 10^{-3}$. We compare the QFI from preparing each initial state to the ultimate precision limit, the extended-channel QFI (ECQFI). The shaded gray region is thus inaccessible [but the upper bound at N is loose in (b)]. In the lossy case, the SMSV and the TMSV with $\eta_A > 0$ exhibit the Rayleigh curse, i.e., QFI $\to 0$ as $\sigma \to 0$, and the Schrödinger's cat state has the same QFI as a coherent state for approximately $N > 3$.

D. Lossy case

Any actual experiment will experience loss that will dramatically change the ECQFI and optimal initial state compared to the lossless case. Consider using the initial state that is optimal in the lossless case, an SMSV state. Given a loss η occurring before the encoding, then the total channel becomes $\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda_\eta^{\mathrm{loss}}$, where we assume that the classical noise is negligible. By Eq. (19), the QFI for an SMSV state in the high-energy limit of $N \to \infty$ is

$$\mathcal{I}_Q(\sigma) = \frac{8\sigma^2}{\left(\eta + 2\sigma^2\right)^2}. \tag{34}$$

In the small-loss limit, $\eta \ll \sigma^2$, the QFI is approximately $2/(\eta + \sigma^2)$, which will be shown to be optimal at any σ^2. In the experimentally relevant loss-dominated regime of $0 < \sigma^2 \ll \eta$, the SMSV performs poorly. The QFI for an SMSV vanishes as $\sigma \to 0$: it suffers from the Rayleigh curse and performs even worse than vacuum, as shown in Fig. 5. This raises the question: What is the ECQFI in the presence of significant loss, $\eta \gg \sigma^2$, and which initial states saturate this bound?

We prove the following claim about the optimal protocol in the lossy case.

Claim 3. The ECQFI given a loss η before the encoding is given by

$$\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma) = \frac{2}{\eta + \sigma^2}, \tag{35}$$

which is attained by preparing a TMSV state with noiseless ancilla ($\eta_A = 0$) in the high-energy limit of $N \to \infty$.

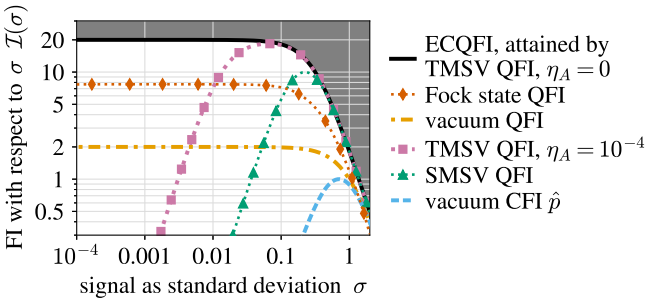

FIG. 5. The FI versus the standard deviation σ for different initial states, indicated in the legend, with a loss of $\eta = 0.1$ occurring before the encoding. The high-energy limit of the ECQFI and the squeezed states is shown. The shaded gray region is inaccessible. For large σ, e.g., $\sigma \sim \eta$, we calculate the QFI for the Fock state $|8\rangle$ numerically using a truncated Hilbert space of dimension 50.

The proof of Claim 3 is given in Appendix C 3, where we use the same method as the lossless case but with different purifications, inspired by the deterministic case [18]. The optimal measurement of a TMSV state with noiseless ancilla is to perform number-resolving measurements of the two modes after first applying antisqueezing and, in general, a beam splitter [37]. Suppose that we instead perform number-resolving measurements directly on the two modes of a TMSV state with average number per mode $n \in \mathbb{Z}$ without first antisqueezing. Then, the CFI does not attain the TMSV QFI and instead equals the lower QFI of the equivalent Fock state $|n\rangle$ given below.

We have also determined the following upper bound on the ECQFI for a given finite $\langle \hat{n} \rangle = N$:

$$\mathcal{I}_Q^{\Lambda'_\sigma, N} < \frac{4}{2(\eta + \sigma^2) + (1-\eta)\xi_N^{-1}}, \tag{36}$$

but this upper bound is not tight for $\eta > 0$ and a fixed finite N. For example, a TMSV state with noiseless ancilla ($\eta_A = 0$) does not saturate this upper bound for finite N, as shown in Fig. 4(b).

For finite $\sigma > 0$, if the loss η_A on the ancilla mode before the encoding is small, $\eta_A \ll \sigma^2, \eta$, then a TMSV state saturates the ECQFI in Eq. (35) in the high-energy limit of $N \to \infty$, as shown in Appendix D. Experimentally, however, this requirement is likely too stringent to probe the small signals of interest. In the realistic regime of $\sigma^2 \ll \eta, \eta_A$, a TMSV state does not saturate the ECQFI.

In the $\sigma \to 0$ limit, all squeezed Gaussian states suffer the Rayleigh curse by Claim 1, provided that loss occurs on every mode (i.e., with fixed $\eta, \eta_A > 0$). The only Gaussian states with nonvanishing QFI are then the coherent states, which still have a QFI of $2/(1+\sigma^2)$, since they remain coherent after the loss before the encoding. This raises the question of whether it is possible to attain the ECQFI in the realistic regime of $\sigma^2 \ll \eta, \eta_A$ by using non-Gaussian states. We address this question in Sec. V F below and show numerically that indeed it appears to be possible.

E. Limit of small signals

Before moving to discuss our numerical results with non-Gaussian states, let us gain some more understanding of the small-signal $\sigma \to 0$ limit. In the lossless case, we have already observed that the QFI in this limit is $4\mathrm{Var}[\hat{x}]$ [51]. Here, we want to understand the lossy case in which the state is mixed before the channel is applied. In Appendix E, we prove the following general claim.

Claim 4. Given an initial state $\hat{\rho}$ and the random unitary channel

$$\Lambda_\sigma(\hat{\rho}) = \int_{-\infty}^{\infty} \mathrm{d}\theta \; p(\theta) \, \hat{U}_\theta \hat{\rho} \hat{U}_\theta^\dagger, \tag{37}$$

where $p(\theta) \sim \mathcal{N}(0, \sigma^2)$ and $\hat{U}_\theta$ is unitary, then the QFI is

$$\mathcal{I}_Q(\sigma = 0) = 4\langle \hat{H} \hat{\Pi}_\perp \hat{H} \rangle,$$

where $\hat{H} = i\hat{U}_0^\dagger \dot{\hat{U}}_0$ is the Hermitian generator at $\sigma = 0$, $\hat{\Pi}_\perp$ is the projection operator onto the null space of $\hat{\rho}$, and the expectation value is calculated with respect to $\hat{\rho}$, i.e., $\langle \hat{O} \rangle = \mathrm{Tr}[\hat{\rho}\hat{O}]$. An optimal measurement that attains the QFI is measurement of $\hat{U}_0^\dagger \hat{\Pi}_\perp \hat{U}_0$.

This result means that $\hat{H}$ needs to map some of $\hat{\rho}$ into its null space to obtain information about σ in the limit of $\sigma \to 0$. In particular, if $\hat{\rho}$ is full rank, then it suffers the Rayleigh curse.

For our random-displacement channel, given a pure initial state $|\psi\rangle$, then $\mathcal{I}_Q(\sigma = 0) = 4\langle \hat{x} \hat{\Pi}_\perp \hat{x} \rangle$, where $\hat{\rho} = \Lambda_\eta^{\mathrm{loss}}(|\psi\rangle\langle\psi|)$ is the state after the loss but before the encoding. This result is useful when $\mathcal{I}_Q(\sigma = 0)$ is nonzero. For example, it implies that, for a Fock state $|N\rangle$, the QFI is $\mathcal{I}_Q(\sigma = 0) = 2(1-\eta)^N(N+1)$ and is attained by number measurement.

What is the optimal initial state in this limit of $\sigma \to 0$? The ECQFI is not well defined unless we specify the order of limits since, by Eq. (29),

$$\lim_{\sigma \to 0} \mathcal{I}_Q^{\Lambda_\theta}(\theta) \neq \sup_{|\Psi\rangle} \lim_{\sigma \to 0} \mathcal{I}_Q^{(\Lambda_\theta \otimes 1_A)(|\Psi\rangle\langle\Psi|)}(\theta).$$

Nevertheless, we claim that $\lim_{\sigma \to 0} \mathcal{I}_Q^{\Lambda_\theta}(\theta)$ is the relevant quantity of interest since, in practice, the signal is small but finite $0 < \sigma \ll 1$ and we can only search for signals above the classical noise floor. In the following subsection, we discuss a family of initial states that are numerically optimal for a fixed small but finite $0 < \sigma \ll 1$.

An open question about this limit of $\sigma \to 0$ is whether the only single-mode pure initial states that are finite rank after the loss $\Lambda_\eta^{\mathrm{loss}}$ are either finite superpositions of coherent states or bounded in the Fock basis.

We discuss a generalization of Claim 4 in Appendix F. Also, in Appendix G, we give an example of how the upper bound on the ECQFI, which is analogous to Eq. (32), can be loose for random unitary channels acting on finite-dimensional systems.

F. Non-Gaussian states

We now explore whether there exist non-Gaussian states of the probe that can outperform the Gaussian states and saturate the ECQFI in Eq. (35) in the relevant high-loss regime $\sigma^2 \ll \eta, \eta_A$.

We start by analyzing the QFI with Fock states. As shown above, the QFI with a Fock state $|N\rangle$ is $\mathcal{I}_Q(\sigma = 0) = 2(1-\eta)^N(N+1)$. The optimal Fock state $|N\rangle$ is thus $N \approx -(1/\log(1-\eta)) - 1$, where the Bose enhancement factor $(N+1)$ balances with the loss factor $(1-\eta)^N$

to achieve a QFI of $\mathcal{I}_Q(\sigma=0)\approx -2/e(1-\eta)\log(1-\eta)$. In the limit of small η, $\mathcal{I}_Q(\sigma=0)\approx 2/e\eta$, which misses the ECQFI by roughly a factor of e, or a penalty of roughly 4.3 dB. This is shown in Fig. 4(b) for $\eta=0.1$, for which the optimal Fock QFI is $\mathcal{I}_Q(\sigma=0)=7.75$ at $N=8,9$, while the ECQFI is $2/\eta=20$. That is, for a loss of 10%, we only need to prepare a Fock state of eight particles to get within 4.1 dB of the ultimate limit. This may be achievable experimentally in the microwave regime as, e.g., Fock states of up to 100 photons have been generated in superconducting microwave cavities [54] and Fock states of four photons have already been used in axion searches [55], as we discuss later. The key takeaway here is that for $\sigma^2\ll\eta,\eta_A$, Fock states outperform all Gaussian states but do not attain the ECQFI in Eq. (35), although they may be a good first step experimentally.

To attain the ECQFI for $0<\sigma^2\ll\eta,\eta_A$, we now consider finite-energy GKP states. We review these states in Appendix H; in particular, the $|\mathrm{GKP}_\Delta\rangle$ family of finite-energy states given by a superposition of displaced SMSV states of width $\propto\Delta$ within a Gaussian window of width $\propto\Delta^{-1}$. To calculate the QFI with respect to σ numerically, we model $|\mathrm{GKP}_\Delta\rangle$ in the Fock basis of a truncated Hilbert space. For the lossless case and $\Delta\ll 1$, the overall variance of the pure initial state $|\mathrm{GKP}_\Delta\rangle$ is $\mathrm{Var}[\hat{x}]=\mathrm{Var}[\hat{p}]=\frac{1}{2}\Delta^{-2}$, which equals the average occupation number $\langle\hat{n}\rangle=N$ [56] such that the QFI for small σ is $4N$, as shown in Fig. 4(a). GKP states are promising for quantum computing and error-correction applications and much engineering is being done to produce them reliably [57]. For example, controlled displacements with a qubit can produce any quantum state of a bosonic mode, at least in principle [58]. Fock states, binomial-code states, and GKP states have been prepared experimentally using this method [59]. To saturate the QFI of the GKP state above, a non-Gaussian measurement is once again required. We numerically determine the Fock-basis coefficients of the optimal set of orthogonal states to project onto for $\sigma>0$. These states are more complicated than number or momentum eigenstates. We do not know how these states relate to the GKP states or what they represent physically, but we suspect that realizing this measurement will be as challenging as producing the GKP states themselves. We defer determining how to realize this measurement to future work.

For the lossy case, e.g., a loss of $\eta=0.1$ occurring before the encoding, the QFI from preparing finite-energy GKP states $|\mathrm{GKP}_\Delta\rangle$ converges numerically to the QFI of a TMSV state with noiseless ancilla ($\eta_A=0$) for $N>10$, i.e., $\Delta<0.2$, as shown in Fig. 4(b). At higher energies, the QFI increases toward the ECQFI. For example, with $\eta=0.1$ and $\sigma=10^{-3}$, $|\mathrm{GKP}_\Delta\rangle$ attains a QFI above 19, within 95% of the ECQFI of 20, using states with $\langle\hat{n}\rangle=N>100$ in a truncated Hilbert space of dimension $\mathcal{O}(1000)$. Note that we have only considered the small-signal σ behavior here, since it is the most relevant regime, and not the large-σ behavior shown in Fig. 5 for the other initial states.

We conjecture that preparing finite-energy GKP states of higher average number, which will require more peaks and a larger truncated Hilbert space, can get arbitrarily close to the ECQFI for any fixed $\sigma^2\ll\eta$. We expect the convergence to be slow given that a TMSV state with noiseless ancilla ($\eta_A=0$) only converges asymptotically. This conjecture is based on the above numerics and, heuristically, the connection between the GKP and TMSV infinite-energy states discussed below. It also would be interesting to understand the performance in the limit of $\sigma\to 0$.

Moreover, we conjecture that the ECQFI for a fixed finite $\langle\hat{n}\rangle=N$ is saturated by preparing a TMSV state with noiseless ancilla for all N. Numerically, we have searched for different non-Gaussian single-mode states that perform better than a TMSV state with noiseless ancilla with the same large $\langle\hat{n}\rangle=N$ but have not found any. We describe our numerical methods in Appendix I. Briefly, we have found that sparse superpositions of finitely many Fock states also outperform Fock states and approach the ECQFI at high energies. For example, we have found a sparse state $|\psi_{\mathrm{num.}}\rangle=\sum_{j=0}^{23}c_j|20j\rangle$ with $\langle\hat{n}\rangle=158.9$ and a QFI of 18.4, within 9% of the ECQFI of 20 for $\sigma^2=10^{-6}$ and $\eta=0.1$. Intuitively, e.g., the signal trajectory from $|0\rangle$ to $|1\rangle$ dominates the loss trajectory from $|20\rangle$ to $|1\rangle$ for finite signals $\sigma^2=10^{-6}$ and $\eta=0.1$. These sparse states are similar to optimized binomial quantum error-correcting codes [60].

Finally, we remark that the high numerical performance of the finite-energy GKP states for sensing a random displacement in the presence of loss is intriguing, as the GKP infinite-energy state was originally designed for the correction of random-displacement noise [52]. Both the GKP infinite-energy state and the TMSV infinite-energy state can be used to form error-correction codes that are sensitive to random displacement signals along $\hat{p}$ yet protected against random-displacement noise along $\hat{x}$. (We discuss this further in Appendix D.) In the next subsection, we will show that this property makes these states resilient to classical noise along $\hat{x}$. It would be interesting to understand how this property is related to their performance in the presence of loss.

G. Classical noise case

We now address the case of significant classical noise such that the total channel is $\Lambda'_\sigma=\Lambda_\sigma\circ\Lambda^{\mathrm{noise}}_{\Sigma_C}$ for a given classical noise matrix Σ_C in Eq. (22). Losses are not present unless otherwise noted. In Appendix J, we prove that the ECQFI and optimal initial state depend on Σ_C as follows.

First, suppose that the classical noise is confined to the same quadrature, $\hat{p}$, as the signal, i.e., $\Sigma_C = \mathrm{diag}\left(0, \sigma_p^2\right)$ with $\sigma_p > 0$. Then, the Rayleigh curse is unavoidable for $\sigma \ll \sigma_p$ and the noiseless ECQFI cannot be recovered, since there is no way to distinguish the signal and the noise. For a fixed finite $\sigma > 0$, the ECQFI is

$$\mathcal{I}_Q^{\Lambda'_\sigma, N}(\sigma) = \frac{4\sigma^2}{(\sigma^2 + \sigma_p^2)\left[2(\sigma^2 + \sigma_p^2) + \xi_N^{-1}\right]}, \tag{38}$$

which reduces to Eq. (23) in the vacuum case ($N = 0$) with $\sigma_x = 0$. The ECQFI in Eq. (38) is attained for a given $\langle \hat{n} \rangle = N$ by preparing the appropriate SMSV state and performing the noiseless optimal measurement—antisqueezing followed by a number-resolving measurement. In the case of parallel classical noise σ_p and loss η, in the high-energy limit the ECQFI becomes

$$\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma) = \frac{2\sigma^2}{(\eta + \sigma^2 + \sigma_p^2)(\sigma^2 + \sigma_p^2)}, \tag{39}$$

such that, in the experimentally relevant regime of $\sigma^2 \ll \sigma_p^2 \ll \eta$, the ECQFI approaches $2\sigma^2/\eta\sigma_p^2$, which exhibits the Rayleigh curse. In comparison, the QFI from preparing an SMSV state in the high-energy limit is

$$\mathcal{I}_Q(\sigma) = \frac{8\sigma^2}{\left[\eta + 2\left(\sigma^2 + \sigma_p^2\right)\right]^2}, \tag{40}$$

which can be attained by performing the above scheme—antisqueezing followed by a number-resolving measurement. It can also be attained by performing a quadrature measurement, but only in the high-energy limit, as otherwise quadrature measurement is suboptimal for an SMSV state at any given finite N. The ratio of the ECQFI in Eq. (39) to the QFI from preparing an SMSV state in Eq. (40) is thus

$$\mathfrak{g} := \frac{\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma)}{\mathcal{I}_Q(\sigma)} = \frac{\left(\eta + 2\varsigma^2\right)^2}{4\varsigma^2\left(\eta + \varsigma^2\right)}, \tag{41}$$

which is a function of η/ς^2, where $\varsigma^2 := \sigma^2 + \sigma_p^2$ and $\mathfrak{g} \geq 1$ by definition of the ECQFI. If $\eta \ll \varsigma^2$, then this ratio $\mathfrak{g}$ approaches one and preparing an SMSV state is close to optimal. In the above experimentally relevant regime of $\sigma^2 \ll \sigma_p^2 \ll \eta$, however, $\eta \gg \varsigma^2$ and $\mathfrak{g}$ approaches $\eta/4\sigma_p^2$, which is large. When $\mathfrak{g}$ is large, i.e., the classical noise is sufficiently small compared to the loss, then preparing a TMSV state with noiseless ancilla or preparing non-Gaussian states can outperform an SMSV state. If the loss on the ancilla can be made smaller than the classical noise on the system, i.e., $\eta_A \ll \sigma_p^2 \ll \eta$, then a TMSV state with lossy ancilla can also outperform an SMSV state. Developing low-loss quantum memories may thus be a viable direction to pursue experimentally as an alternative to non-Gaussian-state preparation, as the required measurement scheme for the TMSV state is simply Gaussian operations (beam-splitter and squeezing unitaries) followed by photodetection.

Second, suppose that the classical noise is confined to the quadrature opposite to the signal, i.e., $\Sigma_C = \mathrm{diag}(\sigma_x^2, 0)$, with $\sigma_x > 0$. The effect of this opposite quadrature noise is fundamentally different from the effect of noise in the same quadrature as the signal; unlike the same quadrature case, noise in the opposite quadrature can be overcome in the limit of large N and does not necessarily lead to a Rayleigh curse. For fixed finite values of $\sigma, \sigma_x > 0$, the noiseless ECQFI in Eq. (32) can be recovered by preparing an SMSV state in the high-energy limit of $N \to \infty$. Intuitively, the classical noise can be squeezed after the encoding while simultaneously antisqueezing the signal. A given SMSV state with fixed finite $\langle \hat{n} \rangle = N$, however, still exhibits the Rayleigh curse as $\sigma \to 0$, such that an SMSV state may not be optimal for a given $\langle \hat{n} \rangle = N$ and σ. In comparison, there exist finite-energy states, e.g., a TMSV state with noiseless ancilla, that do not exhibit the Rayleigh curse and that recover their respective noiseless QFI in the high-energy limit, e.g., $\lim_{N\to\infty} \lim_{\sigma\to 0} \mathcal{I}_Q(\sigma) = 4N$ for a TMSV state.

Finally, suppose that the classical noise appears in both quadratures, i.e., $\Sigma_C = \mathrm{diag}\left(\sigma_x^2, \sigma_p^2\right)$ with $\sigma_x, \sigma_p > 0$. For example, suppose that the classical noise is isotropic with $\Sigma_C = \mathrm{diag}(\sigma_C^2, \sigma_C^2)$ and $\sigma_C > 0$. Then, the Rayleigh curse is unavoidable, as $\sigma \to 0$ with σ_C fixed and the noiseless ECQFI cannot be recovered. The ECQFI with $\sigma_x = 0$ in Eq. (38), however, can be recovered for a fixed finite $\sigma > 0$ by preparing an SMSV state in the high-energy limit despite $\sigma_x > 0$. In the isotropic case, this means simply replacing σ_p by σ_C in Eq. (38).

VI. SIMULTANEOUS ESTIMATION OF THE MEAN AND VARIANCE

We have focused so far on estimating the variance of a random process. In contrast, most work on quantum metrology up to now has been dedicated to estimating the mean value of a signal rather than its variance. Formally, estimating the deterministic encoding of a signal is equivalent to mean estimation in the zero-variance limit. Here, we unify these two efforts as we consider the optimal simultaneous estimation of the mean and the variance. We emphasize that we only consider the case in which the initial state is vacuum such that any loss occurring before the encoding can be ignored.

A. Review of quantum multiparameter estimation

We now introduce the tools of quantum multiparameter estimation, in contrast to single-parameter estimation as reviewed in Sec. II D. In the multiparameter case, a vector of real parameters $\vec{\theta}$ is encoded in a quantum state $\hat{\rho}(\vec{\theta})$. Given M independent and identically distributed measurements of the probability distribution $p(x|\vec{\theta})$, the CCRB provides a tight lower bound on the covariance matrix, $\Sigma_{\vec{\theta}}$, of the estimators of $\vec{\theta}$ as $\Sigma_{\vec{\theta}} \geq (1/M)[\mathcal{I}_C(\vec{\theta})]^{-1}$, where matrix inequalities $A \geq B$ are interpreted henceforth as $A - B$ being positive semidefinite and the classical Fisher information matrix (CFIM) is

$$\mathcal{I}_C(\vec{\theta})_{i,j} = \int_{-\infty}^{\infty} \mathrm{d}x \, \frac{\partial_{\theta_i} p(x|\vec{\theta}) \partial_{\theta_j} p(x|\vec{\theta})}{p(x|\vec{\theta})}.$$

As in the single-parameter case, the CCRB is asymptotically saturable with maximal likelihood estimators and the CFIMs of two independent observations sum: if $p(x|\vec{\theta})$ and $p(y|\vec{\theta})$ are independent distributions, then the total CFIM from observing one outcome from each is the sum of the two CFIMs.

The QCRB provides a further lower bound on the covariance matrix of the estimators given any measurement strategy,

$$\Sigma_{\vec{\theta}} \geq \frac{1}{M}[\mathcal{I}_C(\vec{\theta})]^{-1} \geq \frac{1}{M}[\mathcal{I}_Q(\vec{\theta})]^{-1}.$$

Here, the quantum Fisher information matrix (QFIM) is defined as

$$[\mathcal{I}_Q(\vec{\theta})]_{ij} = \frac{1}{2}\mathrm{Tr}[\hat{\rho}\{\hat{L}_i, \hat{L}_j\}],$$

where the Hermitian operator $\hat{L}_i$ is the symmetric logarithmic derivative (SLD) of $\hat{\rho}$ with respect to θ_i, which is defined implicitly by the equation $\partial_{\theta_i}\hat{\rho} = \frac{1}{2}\{\hat{L}_i, \hat{\rho}\}$. Given the spectral decomposition of $\hat{\rho} = \sum_j p_j |\phi_j\rangle\langle\phi_j|$, the SLDs are given by

$$\hat{L}_i = \sum_{k,l} \frac{2\langle\phi_k|\partial_{\theta_i}\hat{\rho}|\phi_l\rangle}{p_k + p_l} |\phi_k\rangle\langle\phi_l|,$$

where the sum runs over only k, l such that $p_k + p_l > 0$. Similarly to the single-parameter case, the QFIM for a product state $\hat{\rho}_1(\theta) \otimes \hat{\rho}_2(\theta)$ is simply the sum of the individual QFIMs.

Unlike the single-parameter case, however, this bound is not saturable *a priori*. A necessary and sufficient condition for the asymptotic saturability of the QFIM is the weak-commutativity condition [61]

$$\mathrm{Tr}\Big[\hat{\rho}[\hat{L}_i, \hat{L}_j]\Big] = 0, \quad \forall i, j. \tag{42}$$

If this weak-commutativity condition holds, then a joint (i.e., collective) measurement of many copies of $\hat{\rho}$ may be necessary to asymptotically saturate the QFIM. (If this condition does not hold, then the Holevo CRB would instead be required to find a tight bound [17,62].)

B. QCRB for simultaneous estimation

We now consider the problem of estimating μ and σ from the encoding in Eq. (15) but with $p(\alpha) \sim \mathcal{N}(\mu, \sigma^2)$. We assume that the initial state is the vacuum state, such that it becomes the displaced squeezed thermal state with $\vec{\mu} = (0, \mu)$ and $\Sigma = \mathrm{diag}\left(\frac{1}{2}, \frac{1}{2} + \sigma^2\right)$. The QFIM about parameters $\vec{\theta}$ encoded in a single-mode Gaussian state is given by a generalization of Eq. (19) [63]

$$[\mathcal{I}_Q(\vec{\theta})]_{jk} = \frac{\mathrm{Tr}\left[\Sigma^{-1}[\partial_{\theta_j}\Sigma]\Sigma^{-1}[\partial_{\theta_k}\Sigma]\right]}{2(1+\gamma^2)} + \frac{2[\partial_{\theta_j}\gamma][\partial_{\theta_k}\gamma]}{1-\gamma^4} + [\partial_{\theta_j}\vec{\mu}]^{\mathrm{T}}\Sigma^{-1}[\partial_{\theta_k}\vec{\mu}]. \tag{43}$$

Let $\vec{\theta} = (\mu, \sigma)^{\mathrm{T}}$. Then the QFIM for the vacuum state is

$$\mathcal{I}_Q(\mu, \sigma) = \begin{bmatrix} \frac{2}{1+2\sigma^2} & 0 \\ 0 & \frac{2}{1+\sigma^2}, \end{bmatrix} \tag{44}$$

such that the QCRB per independent measurement is $\Delta^2\mu = \Delta^2\sigma = \frac{1}{2}$ for $\sigma \ll 1$. We emphasize here that the M measurements must be independent and identically distributed, which means that the stochastic signal must be uncorrelated between the M measurements, i.e., all off-diagonal components are zero in the covariance matrix of the M experiments in time or space. This assumption will be relevant later when we discuss the application to searching for a stochastic gravitational-wave background.

This bound is the best we could expect for: the estimation variances of σ and μ given by it are the same as in each respective single-parameter case. If the QCRB is tight, therefore, we can estimate the two parameters simultaneously with a precision for each that is identical to its respective single-parameter case.

We check the saturability of the QCRB using the weak-commutativity condition in Eq. (42). By calculating the SLDs for the Gaussian channel [37,64], we find that Eq. (42) holds for the vacuum case. Our remaining task is to determine the optimal asymptotic measurement scheme that saturates the QCRB and whether it must be a joint measurement acting collectively on multiple copies of the state.

C. Separate measurements

For the vacuum case, while a number-resolving measurement of $\hat{n}$ is optimal for estimating σ alone (in the $\mu = 0$ case) and μ alone (in the $\sigma = 0$ case), it is inefficient for estimating the two parameters simultaneously.

Analogously to Eq. (21), the probability of a number-measurement outcome n is

$$p(n) = \int_{-\infty}^{\infty} \mathrm{d}\alpha \ \left|\langle n|\alpha'\rangle\right|^2 p(\alpha)$$
$$= \frac{e^{\frac{-\mu^2}{2(\sigma^2+1)}}\sigma^{2n}}{n!\left(\sigma^2+1\right)^{n+\frac{1}{2}}}\left[\mu^{2n}\frac{(\sigma^{2n}-[4(\sigma^2+1)]^n)}{2^n\sigma^{4n}(\sigma^2+1)^n}\right.$$
$$\left.+n!\mathcal{L}_n^{(-\frac{1}{2})}\left(-\frac{2\mu^2}{4\sigma^2(\sigma^2+1)}\right)\right],$$

where $\mathcal{L}_n^{(\alpha)}(x)$ is the generalized Laguerre polynomial. In the relevant limit of $\mu,\sigma \ll 1$, the probabilities are $p(0) \approx 1-a$ and $p(1) \approx a$, where $a = \frac{1}{2}(\mu^2+\sigma^2)$. These probabilities are degenerate with respect to μ and σ such that only $\sqrt{\mu^2+\sigma^2}$ can be estimated and μ cannot be distinguished from σ. Indeed, the CFIM for a number measurement is

$$\mathcal{I}_C^{\hat{n}}(\mu,\sigma) = \frac{2}{\left(\mu^2+\sigma^2\right)\left(1-\mu^2+\sigma^2\right)}\begin{bmatrix}\mu^2 & \mu\sigma\\ \mu\sigma & \sigma^2,\end{bmatrix} \tag{45}$$

which is singular, and the only nonzero eigenvalue corresponds to the parameter $\sqrt{\mu^2+\sigma^2}$.

For simultaneous estimation of μ and σ, therefore, we need to add a measurement that will break this degeneracy. This can be accomplished by adding a quadrature measurement. In the limit of $\mu,\sigma \ll 1$, quadrature measurement is optimal for estimating μ yet provides no information about σ since it is Rayleigh cursed, as can be observed from the CFIM

$$\mathcal{I}_C^{\hat{p}}(\mu,\sigma) = \begin{bmatrix}\frac{2}{1+2\sigma^2} & 0\\ 0 & \frac{8\sigma^2}{\left(1+2\sigma^2\right)^2}\end{bmatrix} \xrightarrow[\sigma\to 0]{} \begin{pmatrix}2 & 0\\ 0 & 0\end{pmatrix}.$$

Performing a combination of number and quadrature measurements should thus allow for simultaneous estimation. Suppose that we perform many measurements of which the proportion of number measurements is p for some $p \in (0,1)$. For this protocol, the CCRB per measurement is

$$\Sigma = \left[p\mathcal{I}_C^{\hat{n}} + (1-p)\mathcal{I}_C^{\hat{p}}\right]^{-1}$$
$$= \frac{1}{2(1-p)}\begin{pmatrix}1 & -\frac{\mu}{\sigma}\\ -\frac{\mu}{\sigma} & \frac{1-p}{p}+\frac{\mu^2}{p\sigma^2},\end{pmatrix}$$

such that the minimum MSEs are $\Delta^2\mu = 1/2(1-p)$ and $\Delta^2\sigma = 1/2p + \mu^2/2(1-p)p\sigma^2$. The QCRB per measurement of $\Delta^2\mu = \Delta^2\sigma = \frac{1}{2}$ is, therefore, not saturated by this separate measurement strategy.

Intuitively, $\hat{p}$ provides an optimal measurement of μ and $\hat{n}$ provides an optimal measurement of σ for a sufficiently small μ. These observables, however, only weakly commute, since $\langle[\hat{p},\hat{n}]\rangle = -i\langle\hat{x}\rangle = 0$, such that these measurements cannot be performed simultaneously. Since only one of them is performed each time, $\Delta^2\mu$ suffers from a factor-of-$(1-p)$ gap from the QCRB and, similarly, $\Delta^2\sigma$ suffers from a gap of p. Although $\Delta^2\sigma$ suffers an additional uncertainty of $\mu^2/2p(1-p)\sigma^2$, this can be suppressed asymptotically by reducing μ adaptively, i.e., by estimating μ and then nulling by displacing the state back toward the vacuum using the estimate of μ before the number measurement. We describe a suitable adaptive protocol further in Appendix K.

In various cases, the figure of merit is a weighted sum of the variances, i.e., $\mathrm{Tr}[W\Sigma]$, where $W = 2\,\mathrm{diag}(1-w,w)$, with $w \in (0,1)$, is the weight matrix. Informally, the weight w represents how much more we care about estimating σ than μ or vice versa. We thus want to know the optimal proportion of number measurements p for a given w to minimize this figure of merit:

$$\mathrm{Tr}[W\Sigma] = \frac{(-2pw+p+w)\sigma^2+w\mu^2}{(1-p)p\sigma^2}.$$

The optimal proportion can be shown to be

$$p = \frac{w}{w+\sqrt{\frac{w\left[w\mu^2+(1-w)\sigma^2\right]}{\mu^2+\sigma^2}}},$$

such that the optimal figure of merit is

$$\mathrm{Tr}[W\Sigma] = 1+\frac{2\mu^2 w}{\sigma^2}$$
$$+2\frac{\sqrt{\mu^2+\sigma^2}}{\sigma^2}\sqrt{w\left[w\mu^2+(1-w)\sigma^2\right]},$$

which does not saturate the QCRB of $\mathrm{Tr}[W\mathcal{I}_Q^{-1}] = 1$, except in the limit of $w = 0$, which corresponds to estimating μ by solely performing quadrature measurements. This is due to two penalty terms. The first penalty term, $2\mu^2 w/\sigma^2$, stems from nonzero μ and can be asymptotically removed with an adaptive protocol. The second penalty term is fundamental and appears for any $w \neq 0,1$. In the limit $\mu \to 0$, which is attainable with an adaptive protocol, the optimal p is $p = \sqrt{w}/(\sqrt{w}+\sqrt{1-w})$ and the optimal figure of merit is $1+2\sqrt{w(1-w)}$, which attains the QCRB for $w = 0,1$. The worst-case scenario corresponds to equal weights, i.e., $w = 0.5$, where the optimal proportion of number and quadrature measurements is also equal, i.e., $p = 0.5$, and the figure of merit is $\mathrm{Tr}[W\Sigma] = 2+2\mu^2/\sigma^2$, since $\Delta^2\mu = 1$ and $\Delta^2\sigma = 1+2\mu^2/\sigma^2$.

Numerical results of the adaptive and nonadaptive schemes are shown in Fig. 6, where we plot the MSE in

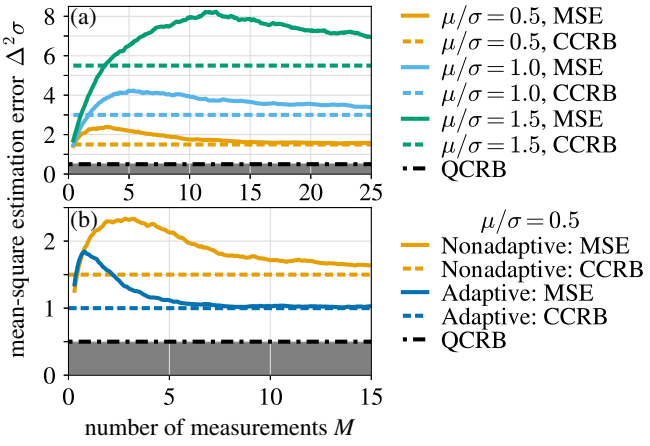

FIG. 6. The numerical mean-square estimation error (MSE) compared to the Cramér-Rao bounds (CRBs) for simultaneous estimation of μ and σ versus the number of independent measurements M. The shaded gray region is inaccessible with the input state fixed as the vacuum state. (a) The nonadaptive-separate-measurements scheme with $\sigma = 0.05$ and different ratios of μ/σ. (b) Adaptive separate measurements are more precise than the nonadaptive strategy with $\sigma = 0.05$ and $\mu/\sigma = 0.5$ but neither reach the QCRB asymptotically as $M \to \infty$.

estimating σ. Denoting the M outcomes of the quadrature measurement and number measurement as $\{p_i\}_{i=1}^M$ and $\{n_i\}_{i=1}^M$ respectively, the estimator of μ is $\tilde{\mu} = (1/M)\sum_i p_i$ and the estimator of σ is $\varsigma = \sqrt{(2/M)\sum_i n_i - \tilde{\mu}^2}$. The distribution of ς is not Gaussian and hence it only converges to the CCRB of $\Delta^2\sigma = 1$ asymptotically, as shown in Fig. 6.

While we have shown that there is no combination of number and quadrature measurements that can saturate the QCRB, we would need to calculate the Nagaoka-Hayashi bound to prove that no other separate measurement scheme can attain the QCRB [65,66].

D. Joint measurements

Suppose that we perform joint measurements on the M independent and identical copies of the final state $\hat{\rho} = \Lambda_\sigma(|0\rangle\langle 0|)$. This collective state $\hat{\rho}^{\otimes M}$ is an M-mode Gaussian state with $2M$-by-1 mean vector $\vec{\mu} = (0, \mu, 0, \mu, \cdots, 0, \mu)^{\mathrm{T}}$ and $2M \times 2M$ covariance matrix $\Sigma = \bigoplus_{j=1}^M \mathrm{diag}\left(\frac{1}{2}, \frac{1}{2} + \sigma^2\right)$ in the basis $(\hat{x}_1, \hat{p}_1, \cdots, \hat{x}_M, \hat{p}_M)^{\mathrm{T}}$. This M-mode state can be transformed to the (anti)symmetric basis such that $\vec{\mu} = (\sqrt{M}\mu, 0, 0, 0, \cdots, 0, 0)^{\mathrm{T}}$, where the symmetric mode is listed first and Σ is unchanged. We emphasize that the elements of this (anti)symmetric basis are orthogonal and commute. This shows that measuring the $\hat{p}$ quadrature of the symmetric mode is a sufficient statistic for μ. The remaining $M-1$ antisymmetric modes can then be used to measure σ performing a number measurement on each mode. Absorbing the M factors from the CRBs, the CFIM for this joint measurement is $2\,\mathrm{diag}(M, M-1)$, which saturates the QFIM in Eq. (44) of $2M\,\mathrm{diag}(1,1)$ in the asymptotic limit of $M \to \infty$.

It is therefore possible to saturate the QFIM if we can perform joint measurements on M independent and identical copies of the final state. These M copies can correspond to different experiments distributed in space or time. For uncorrelated experiments distributed in space, the required transformation to symmetric and antisymmetric modes could be done using a sequence of beam-splitter unitaries; while for uncorrelated experiments distributed in time, the transformation may be done using quantum memories, as further discussed in Sec. VII A.

This protocol is an instance of the following general statement for any quantum system.

Claim 5. Given an initial pure state $|\psi\rangle$ and the random unitary channel

$$\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|) = \int_{-\infty}^{\infty} \mathrm{d}\theta\; p\,(\theta)\; \hat{U}_\theta |\psi\rangle\langle\psi| \hat{U}_\theta^\dagger, \tag{46}$$

where $p(\theta) \sim \mathcal{N}(\mu, \sigma^2)$ and $\hat{U}_\theta = \exp(-i\theta\hat{H})$, then the QFI with respect to either $\mu \to 0$ or $\sigma \to 0$ equals $4\mathrm{Var}[\hat{H}]$ and the QFIM is simultaneously saturable with a joint measurement in the asymptotic limit of $M \to \infty$.

The proof of Claim 5 is given in Appendix L. The key idea is the same as in our particular case above: while the parameter μ displaces the state in the symmetric subspace of the M copies, σ takes the state out of the symmetric subspace. The QFI with respect to μ is saturated by measuring a suitable basis in the symmetric subspace, while the QFI with respect to σ is saturated by measuring the projection onto the antisymmetric subspace. Similar protocols to saturate the QCRB by projecting onto the symmetric and antisymmetric subspaces have previously been proposed for superresolution in imaging [67] and estimating phase diffusion in qubits [68].

Lastly, we remark that preparing collective initial states could improve the joint measurement of the mean and variance in the presence of losses. The optimal collective strategy is to prepare the symmetric mode in an SMSV state to sense the mean μ and to prepare the antisymmetric modes in the optimal non-Gaussian states to sense the standard deviation σ, e.g., GKP states.

VII. STOCHASTIC WAVEFORM ESTIMATION, IMPLEMENTATION, AND APPLICATION

We now solve the continuous-estimation problem of determining $S_{yy}(\Omega)$ at each frequency Ω. By the chain rule from σ to $S_{yy}(\Omega)$ using Eq. (14), the QFI with respect to

$S_{yy}(\Omega)$ is

$$\mathcal{I}_Q[S_{yy}(\Omega)] = \frac{G_{py}(\Omega)^2}{8\sigma^2}\mathcal{I}_Q(\sigma). \tag{47}$$

where we include a factor of 2 to account for adding the information from the independent real and imaginary parts of the mode at Ω.

By Eqs. (39) and (47), the ECQFI for sensing $S_{yy}(\Omega)$ in the presence of an input loss η and signal-referred parallel classical noise $\mathcal{S}^C(\Omega)$ from Eq. (26) is thus equal to

$$\mathcal{I}_Q^{\Lambda'_\sigma}[S_{yy}(\Omega)] = \left[S_{yy}(\Omega) + \mathcal{S}^C(\Omega) + 2\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)\right]^{-1} \cdot [S_{yy}(\Omega) + \mathcal{S}^C(\Omega)]^{-1}, \tag{48}$$

where we have used Eq. (27) to relate the gain $G_{py}(\Omega)$ to the signal-referred quantum shot noise from the vacuum, $\mathcal{S}^{Q,\mathrm{vac}}(\Omega)$. For example, if the initial state is vacuum, then the QFI with respect to $S_{yy}(\Omega)$ is given above by setting $\eta = 1$ and is attained by performing a number-resolving measurement at each frequency. To attain the ECQFI in Eq. (48) at each frequency Ω and increase the QFI beyond the vacuum case by roughly a factor of $1/\eta$ in the vanishing-signal limit, we need to instead prepare non-Gaussian initial states and perform non-Gaussian measurements independently in the $\cos(\Omega t)$ and $\sin(\Omega t)$ signal components of Eq. (10). This is experimentally demanding, as we discuss in detail later.

Let us compare the ECQFI in Eq. (48) to the QFI from preparing an SMSV state. In the limit where the signal and classical noise are small compared to the quantum shot noise, i.e., $S_{yy}(\Omega), \mathcal{S}^C(\Omega) \ll \eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)$, the ECQFI in Eq. (48) may be approximated as

$$\mathcal{I}_Q^{\Lambda'_\sigma}[S_{yy}(\Omega)] \approx \left\{2\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)[S_{yy}(\Omega) + \mathcal{S}^C(\Omega)]\right\}^{-1}.$$

In comparison, by Eqs. (40) and (47), the QFI from preparing an SMSV state is attained in the high-energy limit by performing a quadrature measurement and is given by

$$\mathcal{I}_Q[S_{yy}(\Omega)] = [S_{yy}(\Omega) + \mathcal{S}^C(\Omega) + \eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)]^{-2},$$

which is approximately $\mathcal{I}_Q[S_{yy}(\Omega)] \approx [\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)]^{-2}$ in the shot-noise–dominated limit. (For example, in the vacuum-injection case, the CFI from quadrature measurement is given by setting $\eta = 1$.) The ratio between the ECQFI and the QFI from preparing an SMSV state is thus large and given by [cf. Eq. (41)]

$$\frac{\mathcal{I}_Q^{\Lambda'_\sigma}[S_{yy}(\Omega)]}{\mathcal{I}_Q[S_{yy}(\Omega)]} \approx \frac{\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)}{2\left[S_{yy}(\Omega) + \mathcal{S}^C(\Omega)\right]},$$

such that experiments where the signal and classical noise are small compared to the quantum shot noise can be significantly accelerated by preparing non-Gaussian states compared to preparing an SMSV state.

In principle, there is a continuum of independent parameters $S_{yy}(\Omega)$ to estimate the spectrum but, in practice, we are often only interested in distinguishing spectra from a family $S_{yy}(\Omega|\vec{\theta})$ parametrized by a finite set of real parameters, $\vec{\theta}$. For example, below, in Sec. VII B, we will examine the single-parameter case in which the spectrum $S_{yy}(\Omega|\alpha) = \alpha\Phi(\Omega)$ is solely determined by the unknown scale parameter α, since the morphology $\Phi(\Omega)$ is known. Additionally, given a limited measurement bandwidth $\Delta\Omega$ and a total measurement interval ΔT, it is only possible to make a finite number of measurements $M = \Delta\Omega\Delta T/\pi$. The task then becomes to estimate $\vec{\theta}$ from the measurement outcomes $\{S_{yy}(\Omega_j|\vec{\theta})\}_{j=1}^M$ at the frequencies $\{\Omega_j\}_{j=1}^M$ such that the QFIM is

$$[\mathcal{I}_Q(\vec{\theta})]_{ij} = \sum_{k=1}^{M}\mathcal{I}_Q[S_{yy}(\Omega_k|\vec{\theta})]\frac{\partial S_{yy}(\Omega_k|\vec{\theta})}{\partial\theta_i}\frac{\partial S_{yy}(\Omega_k|\vec{\theta})}{\partial\theta_j}, \tag{49}$$

where, since the modes at each frequency are independent of each other by linearity, the total QFIM is the sum of the individual QFIMs. Note that the optimal measurement at each frequency to attain the QFIM depends on the initial state.

We may generalize the choice of temporal basis. We have so far assumed that we measure the state of the Fourier component at a given frequency, equivalent to measuring the cosine and sine functions in the time domain as given by Eq. (10). We generalize this description, as the Fourier basis is expressed in a formal limit that is not experimentally accessible. Instead, we may measure the state in an arbitrary temporal basis $\{w_j\}_{j=1}^M$ of orthogonal functions such that $\int_{-\infty}^{\infty}\mathrm{d}t\, w_j(t)w_k^*(t) = \delta_{jk}$. The temporal basis can be chosen to fit the given signal model and, in particular, each temporal mode, w_j, can be associated with a parameter of interest, θ_j. The state of each temporal mode is described by the density matrix $\hat{\rho}_k := \iint_{-\infty}^{\infty}\mathrm{d}\tau\mathrm{d}\tau'\, w_k(\tau)\hat{\rho}(\tau,\tau')w_k^*(\tau')$ and the operators that act on this state may be similarly constructed, e.g., the annihilation operator is $\hat{a}_k := \int_{-\infty}^{\infty}\mathrm{d}\tau\, w_k(\tau)\hat{a}(\tau)$ [69]. Here, $\hat{\rho}(\tau,\tau')$ is the density matrix of a single transverse mode of the outgoing bosonic field where the wave parameter $\tau = t - cz$ depends on time t and the distance along the spatial propagation axis z, along which the mode propagates at speed c. For example, this state could represent a paraxial Gaussian beam of light or a wire acting as a transmission line. Note that we still assume that the noise is stationary.

Using a temporal measurement basis is a significant departure from the deterministic displacement case, in which we record the time series from quadrature measurement and can then choose any temporal basis *post*

hoc using classical processing. Here, we instead directly perform, e.g., a number measurement of each temporal mode, which is highly advantageous compared to quadrature measurement for sensing stochastic signals. The measurement comes at the cost, however, of not being able to later process the data classically to study a different temporal basis. Implementations of the independent preparation of non-Gaussian states and performance of non-Gaussian measurements of each temporal mode must be developed for optimal stochastic waveform estimation. We discuss two possible pathways below—resonant filters and cooperative quantum memories.

A. Potential experimental implementations

We first consider the case of preparing the vacuum state and measuring in the Fourier basis. The QFI for an input vacuum state at a given frequency is saturated by number measurement at that frequency. This measurement can be achieved using a resonant filter to extract a band of Fourier modes followed by a number-resolving measurement. For example, in the optical domain, this may be done using resonating filter cavities and low-background photodetectors [24,70]. This implementation, however, can demand a high number of resonant filters. While in principle we only require one filter per parameter θ, in practice we may require additional filters, e.g., to remove background classical noise. Optically, implementing narrow-band filter cavities may be difficult because long round-trip lengths and low scattering loss are required. Furthermore, implementing the measurement in an arbitrary temporal basis or, e.g., the non-Gaussian measurement required for a GKP finite-energy input state may be challenging to do all optically.

A more promising pathway is to instead prepare the initial states in a dedicated ancilla and then couple the states from that ancilla to the appropriate temporal basis of the incoming bosonic mode of the main device. The time reversal of the coupling process can then load the outgoing bosonic mode from the device into an ancilla to be measured. A key example of this protocol would be to use the two-photon Raman transitions of atoms [71] to transmit and receive states in a given temporal basis. To accomplish this in the future with atom-based experimental platforms, we outline the following five requirements:

(R1) preparing non-Gaussian bosonic initial states using the electronic states of atoms or ensembles of atoms
(R2) implementing the optimal non-Gaussian projective measurements using atomic states
(R3) multiplexing such preparation and measurement procedures across many such atomic ensembles
(R4) achieving high cooperativity couplings that minimize transmission loss into and out of these ancillas
(R5) creating long-lived or distantly distributed memories

In general, an optimal stochastic waveform search would demand simultaneous implementation of (R1)–(R4), which have some competing requirements. (R1) and (R2) are feasible with various platforms and schemes, e.g., preparing Fock states with efficient projective readout [72–78] and coherent population trapping [79,80] of atomic states. (R3) is feasible using lattice traps [81] or tweezer arrays [82]. (R4) is feasible using the Purcell effect via cavity enhancement or nanophotonics [83]. Finally, (R5) is required to realize the optimal joint measurement of the associated mean and variance from Sec. VI D using many independent and identical copies of the state, which may correspond to experiments distributed distantly in space (e.g., observatory networks jointly estimating common deterministic and uncorrelated stochastic signals) or in time (e.g., rare events with collective deterministic and uncorrelated stochastic properties). (R5) has been demonstrated in many kinds of systems [84–86] but its integration with the other requirements is an ongoing effort in the field of emerging quantum technologies.

While, in the above discussion, we emphasize optical sensors and atomic memories, microwave systems are similarly promising. They can leverage superconducting cavities and nonlinear junctions [87–91] to form qubits that can efficiently produce non-Gaussian states [92] across multiplexed devices [93] for fundamental physics and sensing applications [94]. The need for low classical noise suggests that the greatest benefit will be for microwave sensors operating in the $\gg$200-MHz range that can search for signals above the thermal black-body radiation of dilution refrigerators [95], i.e., in the $\sigma \gg \sigma_C$ regime [96].

B. Applications

We now consider fundamental-physics applications that highlight the metrological advantages of non-Gaussian-state preparation and measurement for sensing stochastic signals.

1. Quantum gravity

The Verlinde-Zurek theory of quantum gravity predicts large-scale "geontropic" length fluctuations that scale with the holographic surface area of a given volume of spacetime [1,2]. This signal is predicted to manifest as excess noise in the phase quadrature $\hat{p}$ of an optical Michelson interferometer that scales linearly with the arm length. This scenario is described by Eq. (3) with $\theta = \pi/2$, where $S_{yy}(\Omega) = \alpha\Phi(\Omega)$ for some scale factor α to be estimated, and known signal morphology $\Phi(\Omega)$ (e.g., the fiducial spectrum from Ref. [2]) such that $\sigma^2 = \frac{1}{2}\alpha G_{py}(\Omega)\Phi(\Omega)$ by Eq. (14). Previous observations have constrained α to be less than one [4]. The future Gravity from the

Quantum Entanglement of Space-Time (GQuEST) optical interferometry experiment is proposed to further constrain α [3,70] (see also Ref. [5]).

In GQuEST, the classical noise background is projected to be 7 orders of magnitude below the quantum shot noise, $\sigma_C^2 \sim 10^{-7} \ll \frac{1}{2}$, while the optical loss is projected to be $\eta, \eta_A \sim 0.1$, limited by resonating filter cavities outside the interferometer. (Contributions to the optical loss from inside the interferometer are projected to be as low as 10^{-4}.) For signals below the classical noise background, i.e., $\sigma \ll \sigma_C$, this situation is described by Eq. (39), which exhibits the Rayleigh curse. We search instead for signals above the classical noise background, i.e., $\sigma_C \ll \sigma \ll \eta, \eta_A$. In this loss-dominated regime, preparing squeezed states or performing quadrature measurements introduces the Rayleigh curse. By Eqs. (47) and (49), the integrated QFI with respect to α is

$$\mathcal{I}_Q(\alpha) = \sum_{j=1}^{M} \frac{G_{py}(\Omega_j)\Phi(\Omega_j)}{4\alpha} \mathcal{I}_Q(\sigma).$$

For example, if the initial state is a vacuum state, then, by Eq. (48) with $\eta = 1$, the QFI is

$$\mathcal{I}_Q(\alpha) = \sum_{j=1}^{M} \frac{G_{py}(\Omega_j)\Phi(\Omega_j)}{\alpha[2 + \alpha G_{py}(\Omega_j)\Phi(\Omega_j)]}, \tag{50}$$

which can be attained by number measurement at each frequency Ω_j. For the vacuum case, the results from Sec. III thus indicate that photon counting accelerates the accrual of information by $1/4\sigma^2$, compared to quadrature measurement at each frequency. This acceleration reduces the required observing time or number of independent measurements to reach a given confidence level by the same factor. In comparison, if an SMSV state is prepared, for which quadrature measurement is optimal in the high-energy limit, then the acceleration compared to the vacuum case with quadrature measurement is a factor of $1/\eta^2$ by Eq. (34). Thus, photon counting beats squeezing for this application of stochastic signal estimation since $4\sigma^2 \ll \eta$.

The constraint on α can be further improved by preparing non-Gaussian states. The ECQFI in Eq. (35) indicates that if we prepare the optimal states and perform the optimal measurements, then the accrual of information about α could be accelerated by $1/(4\sigma^2\eta)$ compared to the vacuum case with quadrature measurement, or by $1/\eta$ compared to the vacuum case with number measurement. Thus, optimal state preparation and measurement, if implemented through the aforementioned emerging quantum technology, could significantly accelerate this fundamental-physics application.

2. Stochastic gravitational waves

Gravitational-wave searches may also benefit from our results. The global network of gravitational-wave observatories [7] presently consists of the Laser Interferometric Gravitational-wave Observatory (LIGO) [8], Virgo [9], and the Kamioka Gravitational-Wave Observatory (KAGRA) [10]. Each of these observatories operates as a Michelson interferometer with two ports—the common and differential ports. The common port corresponds to the common mode of the interferometer, i.e., the symmetric mode of the two arms, and a large coherent state is injected into it. The differential port corresponds to the differential (antisymmetric) mode of the interferometer. Information about the gravitational wave is encoded in the output of the differential port: a passing gravitational wave causes small displacements along the phase quadrature of the quantum states reflecting from the output differential port. Presently, these observatories are optimized to estimate the mean value (i.e., signal amplitude) of the gravitational wave by injecting highly squeezed states into the differential port and performing quadrature measurements of the output mode [50]. This is the optimal strategy for sensing deterministic displacements, as discussed in Sec. V B. The gravitational-wave strain sensitivity as an amplitude spectral density is proportional to the inverse square root of the FI.

Here, we consider instead operating a gravitational-wave observatory as a variance (i.e., signal-power) sensor for the stochastic signal from a single stochastic source or an incoherent background of many sources [6,97]. The present strategy outlined above, however, suffers the Rayleigh curse when estimating small stochastic signals: LIGO's present loss of $\eta \leq 0.23$ [50] imposes a detection horizon of $\sigma^2 > \eta/2 \approx 0.12$ by Eq. (34), beyond which it suffers the Rayleigh curse. The possible improvement from using the optimal stochastic sensing protocol, which we discuss below, in sensing astrophysical parameters $\vec{\theta}$ for a given signal model can be calculated from Eq. (49). We focus on astrophysical applications at kilohertz frequencies since there, unlike at low frequencies, LIGO's present classical noise is a factor of 4 below the vacuum shot noise [50], i.e., $\sigma_p^2 \sim 0.13$, and noise models predict that it could be made as low as $\sigma_p^2 \sim 0.05$ in the future [98]. The resulting classical noise horizon ($\sigma^2 > \sigma_p^2$) suffered by the optimal protocol is thus up to 2.3 times further than the loss horizon ($\sigma^2 > \eta/2$) suffered by preparing SMSV states [99]. We focus on detecting stochastic signals within this additional quantum-enhanced range.

The optimal stochastic sensing protocol involves preparing non-Gaussian states at the differential port and performing non-Gaussian measurements. (We assume that we still prepare a large coherent state at the common port.) In particular, we consider photon counting directly in the temporal basis of the gravitational-wave signal templates,

$\{w_k\}_{k=1}^{\infty}$, as follows [3]. A compact binary coalescence strain signal $y(t) = h(t)$ may be decomposed as

$$h(t) = h_{\text{det-LF}}(t) + h_{\text{det-HF}}(t) + h_{\text{stoc}}(t),$$

where $h_{\text{det-LF}}(t)$ and $h_{\text{det-HF}}(t)$ are low-frequency (premerger) and high-frequency (merger and post-merger) deterministic terms, respectively, and $h_{\text{stoc}}(t)$ is a stochastic term. This stochastic term has an associated temporal correlation function $\langle h_{\text{stoc}}(t)h_{\text{stoc}}(t')\rangle = H(t,t')$ which is *a priori* nonstationary. If we assume Gaussianity, then, by the Karhunen-Loéve theorem, we can further decompose this stochastic process into a sum of orthonormal temporal modes, $h_{\text{stoc}}(t) = \sum_{k=1}^{\infty} \sigma_k n_k w_k(t)$, where $\{n_k\}_{k=1}^{\infty}$ are independent unit-normal-distributed random variables and $\{\sigma_k\}_{k=1}^{\infty}$ are the parameters of interest. Preparing non-Gaussian states and photon counting in the basis $\{w_k\}_{k=1}^{\infty}$ could accelerate the search for stochastic sources beyond LIGO's present loss horizon discussed above. Note that this is distinct from the "event-stacking" technique [100–102] for aggregating information on the *deterministic* component of post-merger signals, $h_{\text{det-HF}}(t)$, from multiple binary-neutron-star coalescences [103].

It is also possible to operate a gravitational-wave observatory to simultaneously estimate the mean and variance as a test for unmodeled physics. In this scenario, the temporal basis to measure in is chosen beforehand and both the deterministic and stochastic components of the signal in each basis mode are unknown. For example, there are degrees of freedom in our models of binary neutron stars that are unconstrained by current data, and possibly additional unmodeled physics that would make the source either intrinsically stochastic or have unpredictable variations over the astrophysical population. Also, for black-hole events, we could accelerate tests for a stochastic departure from general relativity using a large ensemble of observed binary black-hole events. In either case, by performing joint measurements to estimate σ from the observation of many events with similar μ with a single observatory, we could perform stronger assumption-free tests for unmodeled physics. We assume here that the stochastic component is uncorrelated between the M different events in time.

To observe μ and σ simultaneously and optimally from a single event, we would need a network of M spatially separated observatories. The stochastic component would be uncorrelated between distant observatories because, e.g., the overlap reduction function between the two LIGO sites is vanishing at kilohertz frequencies [97]. As discussed in Sec. VII A, the challenges in this spatial case are in sending, storing, and receiving the quantum states from each observatory to prepare joint states and perform collective measurements. In Sec. VI D, we have assumed asymptotically large M to find the optimal collective measurement. Here, the number of observatories M is instead likely small such that we do not know what the optimal collective measurement is [104]. Our proposed collective measurement, however, still attains the QFI with respect to μ and is a factor of $(M-1)/M$ away from attaining the QFI with respect to σ. Let us compare this to the local strategy in Sec. VI C, which is a factor of 2 away from the QFI of both μ and σ: if $M = 2$ ($M = 3$) [$M = 4$], then our collective measurement is a factor of 2 more sensitive to μ and simultaneously a factor of 1 $\left(\frac{4}{3}\right)$ $\left[\frac{3}{2}\right]$ more sensitive to σ than the local strategy.

3. Axionic dark matter

Axions are a hypothetical wavelike dark-matter candidate that could also solve the strong $\mathcal{CP}$-problem [11,12]. More generally, if axionlike particles exist, then it is predicted that they should weakly interact with photons at the coupling rate $g_{a\gamma\gamma}$. To search for axions, therefore, we want to estimate $g_{a\gamma\gamma}$ at each frequency, where each frequency corresponds to a possible mass for the axion. Many experiments involving a microwave cavity in the presence of a static magnetic field are searching for axions (see, e.g., Refs. [13,14,55,105–107]). Since the coherence time of the axion is predicted to be short, the resulting displacement of the microwave cavity mode would be stochastic and without a preferred phase on longer timescales.

At a given frequency, this transformation is canonically the symmetric case of the additive Gaussian noise channel $\Lambda_{\Sigma_C}^{\text{noise}}$ in Eq. (22), with $\Sigma_C = \text{diag}(\sigma^2, \sigma^2)$ [108]. This is a two-dimensional random-displacement channel with

$$\Lambda_{\sigma}^{\text{2D}}(\hat{\rho}) \approx \hat{\rho} + \sigma^2 \left(\hat{x}\hat{\rho}\hat{x} + \hat{p}\hat{\rho}\hat{p} - \frac{1}{2}\{\hat{x}^2 + \hat{p}^2, \hat{\rho}\} \right). \quad (51)$$

We want to estimate σ since it is proportional to the coupling constant $g_{a\gamma\gamma}$ at the given frequency (i.e., axion mass), because the excess microwave power is proportional to $g_{a\gamma\gamma}^2$ up to the local dark-matter density, magnetic field strength, and cavity parameters [13,14,57,106]. Optimizing the QFI with respect to σ thus maximizes the axion scan rate, since the scan rate is proportional to the FI integrated over all frequencies [108].

Suppose that the initial pure state encounters loss before the encoding such that the total channel is $\Lambda'_{\sigma} = \Lambda_{\sigma}^{\text{2D}} \circ \Lambda_{\eta}^{\text{loss}}$. The ECQFI of this noisy channel with respect to σ, with energy constraint $\langle \hat{n} \rangle = N$, is known to be [108]

$$\mathcal{I}_Q^{\Lambda'_{\sigma},N}(\sigma) = \frac{4\left(\eta + N\left(\eta + 2\sigma^2\right) + \sigma^2\right)}{\left(\eta + \sigma^2\right)\left(N\left(\eta + 2\sigma^2\right) + \sigma^2 + 1\right)} \xrightarrow[N\to\infty]{} \frac{4}{\eta + \sigma^2},$$

which is attained by a TMSV state with noiseless ancilla ($\eta_A = 0$) in the high-energy limit. (This limit is simply twice the one-dimensional case in Claim 3.) If $\eta_A \neq 0$,

however, then we show in Appendix D that the TMSV QFI vanishes for $\sigma^2 \ll \eta_A$. Moreover, by Claim 1, all squeezed states [109,110] are Rayleigh cursed, i.e., the QFI with respect to σ converges to zero in the limit of $\sigma \to 0$, provided that loss occurs on every mode. In general, we expect to operate in this loss-dominated regime experimentally, $\sigma^2 \ll \eta, \eta_A$, since the excess microwave power is small compared to the cavity loss [109]. In this loss-dominated regime, can we prepare non-Gaussian states instead?

Preparing Fock states for axion searches has been previously demonstrated [55] but without awareness of the relation between the optimal Fock state and the ultimate sensing limit. We establish this relation here. By Appendix F 2, the QFI in the limit of $\sigma \to 0$ is

$$\mathcal{I}_Q(\sigma = 0) = 4\left(\langle \hat{x}\hat{\Pi}_\perp \hat{x}\rangle + \langle \hat{p}\hat{\Pi}_\perp \hat{p}\rangle\right), \tag{52}$$

where $\hat{\Pi}_\perp$ projects onto the null space of the state before Λ_σ^{2D}. For a Fock state $|N\rangle$, Eq. (52) implies that the QFI is $\mathcal{I}_Q(\sigma = 0) = 4(1-\eta)^N(N+1)$. The optimal N depends on the loss η as discussed above, e.g., $N = 8$ for $\eta = 0.1$. The optimal Fock state is better than the Gaussian states but is again roughly a factor of e (4.3 dB) away from the ECQFI. For an input vacuum or Fock state, the optimal measurement is photon counting [95,107]. (Note that here, while we include a large loss before the encoding, we assume that no loss occurs after the encoding.)

Numerically, we observe that preparing finite-energy GKP states can achieve a QFI of at least 36, which is at least 90% of the ECQFI of 40 for $\sigma = 10^{-3}$ and $\eta = 0.1$. Similarly to the one-dimensional case, we conjecture that higher-energy finite-energy GKP states can converge to the ECQFI and are the optimal single-mode states for a given large $\langle \hat{n}\rangle = N$ and small $\sigma \ll \eta$. This means that non-Gaussian states beyond Fock states could further accelerate the search for axionic dark matter.

VIII. CONCLUSIONS AND OUTLOOK

We now draw our main conclusions, summarize our contributions to the literature, discuss the experimental practicality of our work, and look to the future of stochastic waveform estimation.

A. Conclusions

We have found the optimal protocol for stochastic waveform estimation using a linear quantum device. We have simplified the problem to the single-variable estimation of the excess noise in each temporal mode. For realistic losses, all Gaussian protocols exhibit the Rayleigh curse and fail to attain the ultimate precision limit. Instead, we have shown numerically that it is optimal to prepare non-Gaussian states such as finite-energy GKP states. For small signals above the classical noise floor, this non-Gaussian protocol outperforms all Gaussian protocols by orders of magnitude. We have also shown that a joint non-Gaussian measurement protocol is optimal for simultaneously estimating the mean and variance of a stochastic signal. Finally, we have demonstrated how our results may be applied to enhance searches for geontropic fluctuations from quantum gravity, stochastic gravitational-wave signals, and axionic dark matter.

B. Summary of results

For ease of reference, we now summarize our technical contributions to the literature in Table I and as follows:

(1) In Claim 1, we have proved a sufficient and necessary condition for overcoming the Rayleigh curse with Gaussian states. This general result applies beyond our specific encoding channel of random displacements and the context of stochastic waveform estimation. In Sec. V D, we have applied this result to prove that all squeezed states with noisy ancilla are cursed for sensing random displacements in the presence of loss on all modes. This establishes that to beat the vacuum state in the limit of vanishing signal, we must prepare non-Gaussian states.
(2) In Claims 2 and 3 we have calculated the ultimate limit, the ECQFI, on sensing random displacements in the lossless and lossy cases, respectively. We have used a purification-based approach given in Appendix C to accomplish this. In the lossless case, we have found that preparing an SMSV state is optimal given a constraint on the average energy per mode of the initial state; whereas, in the lossy case, we have found that preparing a TMSV state with noiseless ancilla is optimal in the high-energy limit.
(3) In Claim 4, we have calculated the QFI and optimal measurement for any random unitary channel acting on a Hilbert space, not just for random displacements of a bosonic mode, in the limit of vanishing signal. We have generalized this result in Appendix F to sensing the weak decay rate from any Lindbladian master equation.
(4) In Sec. V F, we have analyzed the performance of non-Gaussian states. We have proved that preparing Fock states and directly photon counting beats the

TABLE I. Summary of QFIs for different noise channels and initial states for single-parameter estimation of the variance. The QFIM for multiparameter estimation of the mean and variance for vacuum input is given in Eq. (44).

	Initial quantum state		
Noise channel	Vacuum	Gaussian	Non-Gaussian
Lossless	Eq. (20)	Claim 2	Sec. V C
Lossy	Sec. IV B	Claim 1	Claim 3
Classical noise	Eq. (23)	Eqs. (38) and (40)	Eq. (39)

vacuum state but misses the ECQFI by roughly 4.3 dB. We have shown numerically that finite-energy GKP states approach the ECQFI without requiring a noiseless ancilla like a TMSV state. The optimal non-Gaussian state is likely not unique, as we have also found a variational ansatz, sparse superpositions of finitely many Fock states, using numerical methods given in Appendix I. These sparse states are similar to binomial error correcting codes, outperform the Fock states, and approach the ECQFI if the energy of the input state is unconstrained. In Appendix I, we discuss that finding the optimal state is a biconvex-optimization problem and that the alternating convex search, used previously to efficiently find the optimal state, does not work in this case.

(5) In Sec. V G, we have considered sensing random displacements in the presence of classical noise. We have divided the classical noise into its components parallel and perpendicular to the signal. In the presence of parallel classical noise and no loss, we have shown that preparing an SMSV state is optimal. In the additional presence of loss, we have shown that preparing a GKP state instead is optimal. Meanwhile, in the presence of perpendicular classical noise and no loss, we have shown that preparing a GKP state is not only optimal but can recover the ultimate limit from the case with no classical noise. The above results for GKP states also hold for TMSV states but only if we assume noiseless ancilla, i.e., that the ancilla is free from classical noise and loss.

(6) Finally, in Sec. VI, we have considered the case in which the mean is also unknown, such that the mean and variance of the probability distribution behind the random-displacement signal need to be simultaneously estimated. We have considered a vacuum input state and shown that a number-resolving measurement cannot distinguish the mean from the variance of the signal. Instead, we have shown that a joint measurement of many independent copies of the final state is required to saturate the QFIM, while an adaptive protocol using separate measurements remains a factor of 2 below the QFIM. We have proved in Claim 5 that a similar collective measurement saturates the QFIM for any random unitary channel.

C. Experimental practicality

We now address the practicality of our work. As we discuss in Sec. VII A, there are several challenges in preparing the non-Gaussian states and implementing the non-Gaussian measurements that we propose. Thus, these protocols may take some time to be realized. We believe, however, that there are clear paths to addressing these challenges and that the time that it may require to overcome them does not diminish the importance of our results.

Preparing the optimal Fock state and photon counting can outperform all Gaussian states and get us within roughly 4.3 dB of the ultimate sensing limit, the ECQFI, as we have discussed in Sec. V F. For example, the optimal Fock state contains eight photons given a loss of 10%. This is likely achievable for stochastic waveform estimation in the microwave domain. In superconducting microwave cavities, Fock states have already been generated with up to 100 photons [54] and, in axion searches, the benefits of Fock states of four photons have already been shown [55]. This work adds valuable context to the prior experimental results, establishing where their results stand with respect to the ECQFI for stochastic waveform estimation.

To attain the ultimate limit and beat the Fock states by up to roughly 4.3 dB, we need to prepare more complicated non-Gaussian states such as GKP states or sparse superpositions of finitely many Fock states (which are similar to binomial-code states). Progress to this end is also being made experimentally in the microwave regime, motivated by using these states in quantum computing and error-correction applications. As we have discussed in Sec. V F, GKP states and binomial-code states have been prepared experimentally in a superconducting microwave cavity using controlled displacements with a qubit [59]. Alternatively, as we have discussed in Sec. IV C, it may be possible to reach the ultimate limit by preparing TMSV states and performing Gaussian operators followed by photodetection provided that the loss on the ancilla can be engineered to be sufficiently small compared to the classical noise and the signal.

Finally, we wish to make an analogy to the development of squeezing for gravitational-wave observatories. Four decades have passed from when squeezing was first proposed for LIGO [111] to today, where we measure 5 dB of frequency-dependent squeezing [112]. It may take decades more to produce the 10-dB broadband noise reduction stated for the third generation of gravitational-wave observatories [113,114]. The original squeezing proposal helped catalyze the development of LIGO and gravitational-wave astronomy and led to the maturity of squeezing today. Stochastic waveform estimation provides alternative and parallel science cases to deterministic waveform estimation that are relevant to both particle physics and astrophysics. This work has established the goals and benchmarks to gauge past and future demonstrations of stochastic waveform estimation. This is a valuable contribution given the statistical subtleties in sensing random signals and in systematically defining background noise and degradation processes. While the full extent of our results may not be immediately realized experimentally, many of the preliminary state-preparation and readout technologies have been demonstrated and

our results motivate and provide context to their future development. Our work informs the design of future experiments to maximally benefit from state preparation and optimally extract information.

D. Outlook

There are many related open questions to consider in the future. While we have considered the asymptotic limit and maximum likelihood estimation, it remains to be understood whether we can obtain faster convergence to the fundamental limit or better performance when restricted to a small number of measurements. It would also be interesting to understand the relation with the Bayesian-estimation problem in which, using the notation from Eq. (15), α is instead being estimated and $p(\alpha)$ is its prior distribution. We note that there is no immediate relation between this Bayesian problem and the problem of estimating σ that we consider. For example, in the vacuum case with a Gaussian prior [115,116], quadrature measurement $\hat{p}$ remains asymptotically optimal for estimating deterministic displacements, whereas here $\mathcal{I}_C^{\hat{p}}(\sigma = 0) = 0$. It remains to be understood whether a subtler connection exists between this Bayesian problem and our estimation problem.

Further open questions include the impact of different noise channels, determining the optimal states for fixed finite $\langle \hat{n} \rangle = N$, and the complete estimation of a non-Gaussian stochastic waveform since, e.g., estimating the fourth-order cumulant of a quantum state is of interest to testing nonclassicality [117] and quantum gravity [118–120].

ACKNOWLEDGMENTS

We thank the following people for their advice provided during this research: Rana Adhikari, Evan Hall, Thakur Giriraj Hiranandani, Konrad Lehnert, Katarzyna Macieszczak, Ian MacMillan, Patrick Meyers, Haixing Miao, Swadha Pandey, John Teufel, Mankei Tsang, Sander Vermeulen, Chris Whittle, and Sisi Zhou. We also thank the Caltech Chen Quantum Group and the ANU CGA Squeezer Group. In Fig. 1, we use component graphics from Ref. [121], with permission. The computations presented here were conducted in the Resnick High Performance Computing Center, a facility supported by Resnick Sustainability Institute at the California Institute of Technology. This research is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (Project No. CE170100004). J.W.G. and this research are supported by an Australian Government Research Training Program (RTP) Scholarship and also partially supported by U.S. National Science Foundation (NSF) Grant No. PHY-2011968. In addition, Y.C. acknowledges the support by the Simons Foundation (Award No. 568762) and the NSF Grant No. PHY-2309231. T.G. acknowledges funding provided by the Institute for Quantum Information and Matter and the Quantum Science and Technology Scholarship of the Israel Council for Higher Education. S.A.H. acknowledges support through Australian Research Council Future Fellowship Grant No. FT210100809. J.P. acknowledges support from the U.S. Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research (Grants No. DE-NA0003525 and No. DE-SC0020290), the U.S. DOE, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator, and the NSF (Grant No. PHY-1733907). The Institute for Quantum Information and Matter is an NSF Physics Frontiers Center. L.M.'s photon-counting effort on the GQuEST project is funded in part by the Heising-Simons Foundation through Grant No. 2022-3341. This paper has been assigned LIGO Document No. P2400069.

DATA AVAILABILITY

The data that supports our results uses *Mathematica* [122] and PYTHON [123–131] and is openly available online [132].

APPENDIX A: PROOF OF CLAIM 1

Here, we prove Claim 1 by combining a known result about the eigenvalues of the density matrix and the Williamson decomposition of a Gaussian state [133].

Since $\partial_\sigma \hat{\rho} \to 0$ as $\sigma \to 0$, the optimal measurement basis in this limit is the eigenbasis of the density matrix $\hat{\rho}$ such that the QFI is nonvanishing if and only if there exists an eigenvalue of $\hat{\rho}$ proportional to σ^2 [40]. We therefore need to find a condition on the symplectic eigenvalues of Σ that is equivalent to this condition on the eigenvalues of $\hat{\rho}$. For an M-mode Gaussian state, the Williamson decomposition has the eigenvalues $\prod_{i=1}^{M} \lambda_{k_i}$ and eigenstates $\{\bigotimes_{i=1}^{M} \hat{S}(r_i)\,|k_i\rangle\}_{k_1 \cdots k_M}$, where $\hat{S}(r)$ is the single-mode squeezing operator and $|k_i\rangle$ is a Fock state (potentially of a combination of the original modes, e.g., a TMSV state). The *symplectic eigenvalues* associated with the ith mode are $\frac{1}{2} + \bar{n}_i$, such that the corresponding eigenvalues are $\lambda_{k_i} = (1/(1+\bar{n}_i))(\bar{n}_i/(1+\bar{n}_i))^{k_i}$ with degeneracy 2. Since the QFI is not Rayleigh cursed if and only if there exists some k_i such that $\lambda_{k_i} \propto \sigma^2$ [40], therefore, the QFI is nonvanishing if and only if there exists some mode i such that $\bar{n}_i \propto \sigma^2$ or, equivalently, a symplectic eigenvalue equal to $\frac{1}{2} + k\sigma^2$ for some constant k (not to be confused with k_i).

We remark that for a single-mode Gaussian state, this result can be seen directly from the QFI expression: if $\partial_\sigma \Sigma \to 0$, then the first term in Eq. (19) also vanishes. The second term in Eq. (19), however, is nonzero if and only if $\gamma = 1 - \beta\sigma^2$, which is equivalent to the above condition on the symplectic eigenvalues.

Here, we have assumed that the noisy channel is analytic at $\sigma = 0$ such that we may expand it in a Taylor series. This will usually be the case. For nonanalytic noisy channels, however, this result can be extended: the QFI is nonvanishing if and only if there exists a symplectic eigenvalue equal to $\frac{1}{2} + k\sigma^l$ with $1 < l \leq 2$ [40].

APPENDIX B: OTHER NOISE CHANNELS

We have focused on the case of a noise channel with known parameters occurring before the encoding channel. In this appendix, we briefly discuss noise channels occurring after the encoding channel and explore an example of what happens when the parameters of the noise channel are unknown.

1. Noise occurring after the encoding

Noise occurring after the encoding can be divided into two types: (1) noise associated with or "baked into" the acquisition of the signal or (2) noise associated with the later measurement. In the former case, it is not possible to insert an arbitrary parameter-independent "control" channel between the encoding and noise channels to mitigate the noise. In the latter case of "measurement noise," however, it is possible to entirely overcome the noise by implementing a suitable control channel between the encoding and measurement noise channels.

Let us illustrate this point that measurement noise can be overcome for sensing σ in the case of preparing an SMSV initial state with $\langle \hat{n} \rangle = N$. Note that in the noiseless case, the optimal measurement after the encoding channel is to first antisqueeze and then perform a number-resolving measurement. Here, we assume that we still antisqueeze but then perform a different, fixed, POVM that can be decomposed into a measurement noise channel followed by a number-resolving projective measurement. In, e.g., an optical system, our measurement apparatus is a photodetector and we consider two relevant noise models: detection loss and dark counts. Photodetection with detection loss is modeled as a loss channel $\Lambda^{\text{loss}}_{\eta_{\text{meas}}}$ followed by a projective number measurement. Photodetection with dark counts is modeled as the noise channel $\Lambda^{\text{dark}}_{N_{\text{meas}}}$ followed by a projective number measurement, where $\Lambda^{\text{dark}}_{N_{\text{meas}}}$ is defined as

$$\Lambda^{\text{dark}}_{N_{\text{meas}}}(\hat{\rho}) = \sum_{n=0}^{\infty} \hat{K}_n \hat{\rho} \hat{K}_n^\dagger, \quad \hat{K}_n = \sum_{m=0}^{\infty} \sqrt{p_{\text{th}}(m)} |n+m\rangle\langle n|,$$

where $p_{\text{th}}(m) = (1/(1+N_{\text{meas}}))\,(N_{\text{meas}}/(1+N_{\text{meas}}))^m$ is the number distribution of a thermal state with average number N_{meas}. The results of measuring $\hat{n}$ on $\Lambda^{\text{dark}}_{N_{\text{meas}}}(\hat{\rho})$ are equivalent to convolving the results of measuring $\hat{n}$ on $\hat{\rho}$ with p_{th}. Intuitively, $\Lambda^{\text{dark}}_{N_{\text{meas}}}$ adds m particles to the quantum state with probability $p_{\text{th}}(m)$ independent of the state. In comparison, it is unsuitable to model dark counts as an isotropic classical noise channel $\Lambda^{\text{noise}}_{\Sigma_{C,\text{meas}}}$ with $\Sigma_{C,\text{meas}} = \text{diag}\big(\sigma_{C,\text{meas}}, \sigma_{C,\text{meas}}\big)$, since $\Lambda^{\text{noise}}_{\Sigma_{C,\text{meas}}}$ both adds and subtracts particles at a rate that depends on the quantum state due to Bose enhancement.

We first consider the impact of measurement noise in the absence of any additional control channels. In the noiseless case, given an initial SMSV state, the QFI is $\mathcal{I}_Q(\sigma) = 4\xi_N$ in the limit of $\sigma \ll 1$. In this limit, almost all of the CFI from the projective number measurement after antisqueezing comes from the $\hat{n} = 1$ single-particle detection probability, $p\,(1) \approx \xi_N \sigma^2$. In the case of measurement loss η_{meas}, e.g., detection loss for an optical system, the total channel before the projective number measurement is $\Lambda'_\sigma = \Lambda^{\text{loss}}_{\eta_{\text{meas}}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$, where Λ_{antisqz} is the antisqueezing unitary channel for the given N. The single-particle detection probability is now $p\,(1) \approx (1-\eta_{\text{meas}})\,\xi_N \sigma^2$, such that the CFI from the subsequent projective number measurement falls to $\mathcal{I}^{\hat{n}}_C(\sigma) = (1-\eta_{\text{meas}})\,4\xi_N$ but the Rayleigh curse is not introduced. This is unlike loss occurring before the antisqueezing operation, which does introduce the Rayleigh curse. In contrast, the Rayleigh curse will arise in the case of the dark-count-measurement noise channel. In this case, the total channel before the projective number measurement is $\Lambda'_\sigma = \Lambda^{\text{dark}}_{N_{\text{meas}}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$. The convolved probability distribution of number measurements is then $p(0) = (1-\xi_N\sigma^2)/(1+N_{\text{meas}})$ and $p(n) = N_{\text{meas}}^{n-1}(N_{\text{meas}} + \xi_N\sigma^2)/(1+N_{\text{meas}})^{n+1}$ for $n \geq 1$ such that the CFI is $\mathcal{I}^{\hat{n}}_C(\sigma) = 4\xi_N^2\sigma^2/\big(1-\xi_N\sigma^2\big)\big(\xi_N\sigma^2 + N_{\text{meas}}\big)$, which vanishes in the limit of $\sigma \to 0$ for fixed $\xi_N, N_{\text{meas}} > 0$.

We now show that there exists, in theory, a unitary that allows us to recover the noiseless QFI. After the antisqueezing operation, applying any unitary channel Λ_{swap} that swaps the Fock states $|1\rangle$ and $|k\rangle$ for some $k \gg 1$ and stabilizes the vacuum $|0\rangle$ (up to a phase) will recover the noiseless QFI in the limit of large enough k as proved below. Intuitively, this control unitary channel Λ_{swap} amplifies the signal to make it more tolerant to loss and distinguish it from dark counts. In the case of measurement loss, e.g., detection loss for an optical system, the total channel before the projective number measurement is now $\Lambda'_\sigma = \Lambda^{\text{loss}}_{\eta_{\text{meas}}} \circ \Lambda_{\text{swap}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$. The probability of not detecting zero particles is now $\big(1-\eta^k_{\text{meas}}\big)\,\xi_N\sigma^2$, keeping only the relevant σ^2 terms in the limit of $\sigma \ll 1$, such that the CFI from the subsequent projective number measurement is $\mathcal{I}^{\hat{n}}_C(\sigma) = 4\xi_N\,\big(1-\eta^k_{\text{meas}}\big)$, which recovers the noiseless QFI of $4\xi_N$ in the limit of $k \to \infty$. In the case of the dark-count-measurement noise, the total channel before the projective number measurement is now $\Lambda'_\sigma = \Lambda^{\text{dark}}_{N_{\text{meas}}} \circ \Lambda_{\text{swap}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$. The probability of detecting k or more particles after this channel is

$$p(\hat{n} \geq k) = \xi_N\sigma^2 \left[1 - \left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k\right] + \left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k$$

in the limit of $\xi_N\sigma^2 \ll 1$. The CFI from detecting k or more particles is then

$$\mathcal{I}_C^{\hat{n}} = \frac{4\xi_N^2\sigma^2\left[1-\left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k\right]^2}{\xi_N\sigma^2\left[1-\left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k\right]+\left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k},$$

for fixed $\sigma, N, \bar{n} > 0$. Since, for fixed $\xi_N, \sigma, N_{\text{meas}}$ and $\varepsilon > 0$, $\exists K$ such that $(N_{\text{meas}}/(N_{\text{meas}}+1))^k < \varepsilon$ for $k > K$, the CFI from number measurement recovers the noiseless QFI of $4\xi_N$ in the limit of $k \to \infty$. We defer to future work to find possible ways to implement suitable control unitary channels such as Λ_{swap}.

While we have assumed a specific fixed POVM above, e.g., a noisy photodetector, the results hold more generally [134–136]. Any fixed POVM that is modeled as a loss or dark-counts channel followed by a projective measurement will have the above limitations. For any such POVM, there exists an analogous control unitary channel that can overcome the measurement noise, where the particular control unitary channel needed may depend on the POVM [135,136].

In the case of a deterministic displacement channel, it is known that a phase-sensitive amplifier after the encoding can mitigate subsequent measurement loss in the high-gain limit. This is sometimes called a "Caves's amplifier" in the context of gravitational-wave observatories [111]. In our case, however, a phase-sensitive amplifier after the encoding would introduce the Rayleigh curse and should be avoided.

2. Unknown loss

Here, we discuss the implications of not precisely knowing the loss η that the state $\hat{\rho}$ experiences via the loss channel $\Lambda_\eta^{\text{loss}}$ introduced in Sec. IV B. We illustrate this for the input vacuum case and an indeterminate loss occurring after the encoding but before the measurement.

Suppose that the loss η follows some probability distribution $p(\eta)$. By the central limit theorem, asymptotically, this distribution approaches $p(\eta) \sim \mathcal{N}(\mu_\eta, \sigma_\eta^2)$ with some mean μ_η and standard deviation σ_η. The state after this indeterminate loss is

$$\Lambda_{\mu_\eta,\sigma_\eta}^{\text{indet.}}(\hat{\rho}) = \int_0^1 \mathrm{d}\eta\ \Lambda_\eta^{\text{loss}}(\hat{\rho})p(\eta).$$

We want to know whether this will introduce the Rayleigh curse in the relevant regime of $\sigma_\eta \ll 1$ where the loss is not precisely known but is well constrained.

The nth moment of a weighted average of distributions is the weighted average of their nth moments. The first moment is zero for the vacuum case. The second moment after $\Lambda_\eta^{\text{loss}}$ is linear in η such that the second moment after $\Lambda_{\mu_\eta,\sigma_\eta}^{\text{indet.}}$ equals that after $\Lambda_{\mu_\eta}^{\text{loss}}$, e.g., the signal term is $(1-\mu_\eta)\sigma^2$. The loss indeterminacy σ_η does not affect the first two moments, so, it must be a non-Gaussian perturbation of the third, fourth, and higher-order moments.

For the vacuum case, the final state after $\Lambda_{\mu_\eta,\sigma_\eta}^{\text{indet.}} \circ \Lambda_\sigma$ has the second moment $\langle\hat{p}^2\rangle = \varsigma^2$ and fourth moment $\langle\hat{p}^4\rangle = 3\varsigma^4 + 3\sigma^2\sigma_\eta^2$, where $\varsigma^2 := \frac{1}{2} + (1-\mu_\eta)\sigma^2$. The fourth-order cumulant $3\sigma^2\sigma_\eta^2$ indicates that the distribution has slightly fatter tails than a Gaussian. Numerically, the CFI from number measurement shows that this perturbation is negligible when μ_η is known and σ_η is small. For example, an indeterminacy of $\sigma_\eta/\mu_\eta = 10\%$ leads to an $\mathcal{O}(10^{-4})$ fractional change in the CFI with respect to σ. By convexity, this implies that the QFI also changes negligibly. If we also do not know μ_η, then the CFIM with respect to μ_η and σ is not singular and thus they can be estimated simultaneously.

3. Coarse-grained number-resolving measurements

Realistic measurements are unable to resolve infinitely many different particle numbers. We now consider a coarse-grained number-resolving measurement that can only distinguish between zero counts ($|0\rangle\langle 0|$) and nonzero counts ($\hat{I} - |0\rangle\langle 0|$), but cannot distinguish between, e.g., one or two particles. Suppose that the initial state is vacuum and there is no classical noise. Then, by Eq. (21), the probabilities of these two outcomes are $p(n=0) = 1/\sqrt{\sigma^2+1}$ and $p(n \geq 1) = 1 - p(n=0)$. By Eq. (18), the CFI is thus

$$\mathcal{I}_C(\sigma) = \frac{1}{\left(\sigma^2+1\right)^{5/2}} + \frac{1}{\left(\sigma^2+1\right)^2},$$

which achieves the QFI in Eq. (20) in the vanishing-signal limit of $\sigma \to 0$ but not generally for fixed $\sigma > 0$. This is because all of the information in the vanishing-signal limit comes from the single-particle probability $p(n=1)$, as can be seen from the summand of the CFI in Eq. (18),

$$\frac{[\partial_\sigma p(n)]^2}{p(n)} = \frac{(2n)!\,\sigma^{2n-2}\left(\sigma^2-2n\right)^2}{4^n(n!)^2\left(\sigma^2+1\right)^{n+\frac{5}{2}}},$$

which vanishes in the limit of $\sigma \to 0$ for $n \neq 1$ but saturates the QFI for $n = 1$.

Similarly, for a Fock state $|N\rangle$, all of the information comes from the $p(n=N+1)$ probability in the vanishing-signal limit. For example, this means that for a loss of $\eta = 0.1$, if we prepare the optimal $|8\rangle$, then we need to at least distinguish nine or more particles from eight or fewer particles to have nonzero sensitivity. This is more experimentally difficult than achieving the coarse-grained zero-or-nonzero measurement above. Alternatively, we could

first perform a unitary that maps, e.g., $|N\rangle \mapsto |0\rangle$ and $|N+1\rangle \mapsto |1\rangle$ and then use the coarse-grained zero-or-nonzero measurement above, although implementing such a unitary may be challenging.

More generally, the fact that coarse-grained binary measurements, i.e., those with only two outcomes, suffice to saturate the QFI in the vanishing-signal limit appears to be a general property of Rayleigh-cursed problems. As shown in Refs. [38,40], all the information about σ is encoded in eigenspaces of $\hat{\rho}(\sigma)$, the eigenvalues of which go as σ^2. Hence, the QFI is saturated by a binary measurement that differentiates between $\hat{\rho}(\sigma = 0)$ and the orthogonal eigenspace(s) that go as σ^2. In particular, Claim 1 implies that for Rayleigh-cursed Gaussian states, the optimal measurement corresponds to a symplectic operation followed by a coarse-grained number-resolving measurement. For example, the optimal measurement for a TMSV state corresponds to a symplectic operation that diagonalizes $\hat{\rho}(\sigma = 0)$ followed by a coarse-grained number-resolving measurement of one of the modes.

However, for more complex scenarios such as GKP states and for finite values of σ, binary measurements may not be sufficient and the effect of a realistic counting model is unknown. As we have explained in Sec. V F, we do not know if or how the required measurement for GKP states relates to the process to generate them. This means that we do not know whether a number-resolving measurement is required to implement the measurement for the GKP state, let alone what the effects of the realistic model of counting might be. We leave this to future work once the GKP measurement is first understood.

APPENDIX C: PURIFICATIONS AND THE ECQFI

To prove Claims 2 and 3 about the ECQFI, we first review the established method that we use to optimize over the initial states.

1. Review of purifications

Consider an initial pure state $|\psi\rangle$ and a nonunitary channel Λ_θ such that the final state is $\hat{\rho}(\theta) = \Lambda_\theta(|\psi\rangle\langle\psi|)$. The channel Λ_θ can be purified (also called dilated) to a unitary process $\hat{U}_\theta$ acting on $|\psi\rangle \otimes |\varphi\rangle$ such that the final state is $\hat{\rho}(\theta) = \mathrm{Tr}_A[\hat{U}_\theta(|\psi\rangle \otimes |\varphi\rangle)]$ for all $|\psi\rangle$ and θ. Note that the ancilla $|\varphi\rangle$ is independent of $|\psi\rangle$ and θ but depends on the purification $\hat{U}_\theta$ chosen. The purification of Λ_θ to $\hat{U}_\theta$ should not be confused with the purification of a mixed state or with the ECQFI discussed below.

The choice of purification, however, is not unique. By Uhlmann's theorem for the quantum fidelity, the QFI for a fixed initial state $|\psi\rangle$ is the infimum of the QFI over all possible purifications $\hat{U}_\theta$ of Λ_θ [43,44,108,137]

$$\mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta) = \inf_{\hat{U}_\theta} \mathcal{I}_Q^{\hat{U}_\theta(|\psi\rangle\otimes|\varphi\rangle)}(\theta). \tag{C1}$$

Here, the QFI of the unitary process $\hat{U}_\theta$ is $4\mathrm{Var}[\hat{H}]$ for all θ, where $\hat{H} = -i\hat{U}_\theta^\dagger \dot{\hat{U}}_\theta$ is the Hermitian generator of local displacements in θ and the variance is calculated with respect to the pure initial state $|\psi\rangle \otimes |\varphi\rangle$ [138]. Using this fact, then Eq. (C1) becomes

$$\mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta) = 4 \inf_{\hat{U}_\theta} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}],$$

where $|\varphi\rangle$ and $\hat{H}$ are determined by the purification $\hat{U}_\theta$ of Λ_θ.

The CQFI of Λ_θ in Eq. (28), which optimizes over the initial state, is then given by

$$\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta) = 4 \sup_{|\psi\rangle} \inf_{\hat{U}_\theta} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}].$$

Exchanging the order of maximization and minimization above results in an upper bound on the CQFI of Λ_θ. Remarkably, in Ref. [139] it has been shown that this upper bound is exactly the ECQFI of Λ_θ in Eq. (29); hence

$$\mathcal{I}_Q^{\Lambda_\theta}(\theta) = 4 \inf_{\hat{U}_\theta} \sup_{|\psi\rangle} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}]. \tag{C2}$$

Thus, to attain the ECQFI, it suffices to find a purification $\hat{U}_\theta$ for which the following upper bound is tight:

$$\mathcal{I}_Q^{\Lambda_\theta}(\theta) \le 4 \sup_{|\psi\rangle} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}]. \tag{C3}$$

To find such a purification, we will use the following technique of introducing a "hiding" unitary $\hat{U}_\text{hide}$ [18]. Suppose that we have a purification $\hat{U}_\text{enc.}$ of Λ_θ. We can generate a family of related purifications $\hat{U}_\theta = \hat{U}_\text{hide}\hat{U}_\text{enc.}$ by performing any θ-dependent unitary transformation $\hat{U}_\text{hide}$ on the ancilla, since the ancilla is traced out to recover $\hat{\rho}(\theta)$. Intuitively, this hiding unitary $\hat{U}_\text{hide}$ is meant to remove (or hide) the excess information about θ present in the ancilla after $\hat{U}_\text{enc.}$ to make Eq. (C3) tight.

2. Proof of Claim 2

Here, we prove Claim 2 about the lossless case by using the above purification method with a hiding unitary.

We purify the channel Λ_σ by introducing an ancillary mode 2 prepared in a vacuum state, such that the resulting encoding unitary is a conditional displacement $\hat{U}_\text{enc.} = \exp(-i\sigma\hat{H}_\text{enc.})$ with the generator $\hat{H}_\text{enc.} = \sqrt{2}\hat{x}_1\hat{x}_2$. We also include a hiding unitary $\hat{U}_\text{hide} = \exp(-i\sigma\hat{H}_\text{hide})$ such that the overall purification is $\hat{U}_\sigma = \hat{U}_\text{hide}\hat{U}_\text{enc.}$. Here, $\hat{H}_\text{hide} = \frac{1}{2}g\left(\hat{x}_2\hat{p}_2 + \hat{p}_2\hat{x}_2\right)$ is a squeezing Hamiltonian and g is a real parameter over which to minimize to obtain a tight upper bound in Eq. (C3).

The Hermitian generator for the overall purification $\hat{U}_\sigma$ is

$$\begin{aligned}\hat{H} &= \hat{H}_{\text{enc.}} + \hat{U}_{\text{enc.}}^\dagger \hat{H}_{\text{hide}} \hat{U}_{\text{enc.}}. \\ &= \left(\sqrt{2} + \sqrt{2} g\sigma\right) \hat{x}_1 \hat{x}_2 + \frac{g}{2}\left(\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2\right).\end{aligned}$$

The QFI of $\hat{U}_\sigma |\psi\rangle \otimes |\varphi\rangle$ is therefore

$$4\text{Var}[\hat{H}] = 2\left(\sqrt{2} + \sqrt{2} g\sigma\right)^2 \langle \hat{x}_1^2 \rangle + 2g^2, \tag{C4}$$

where we have assumed that $\langle \hat{x}_1 \rangle = 0$ without loss of generality. Here, since mode 2 is in a vacuum state, we have also used the fact that

$$\text{Var}\left[\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2\right] = 2. \tag{C5}$$

Minimizing the QFI in Eq. (C4) over g, we derive the upper bound to the ECQFI in Eq. (33). This equals the QFI from preparing an SMSV state by Eq. (19).

3. Proof of Claim 3

Here, we prove Claim 3 about the lossy case by using a similar purification method to the lossless case above. Our choice of hiding unitary to handle the loss channel was inspired by the proof of Eq. (31) for the deterministic case [18].

We purify the total noisy channel $\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda_\eta^{\text{loss}}$ as follows. The loss channel $\Lambda_\eta^{\text{loss}}$ is purified to the following beam-splitter unitary between mode 1 and some ancillary mode 3 that starts in vacuum:

$$\begin{aligned}\hat{U}_\eta &= \exp[i \arcsin(\sqrt{\eta})(\hat{a}_1^\dagger \hat{a}_3 + \hat{a}_3^\dagger \hat{a}_1)] \\ &= \exp[i \arcsin(\sqrt{\eta})(\hat{x}_1 \hat{x}_3 + \hat{p}_1 \hat{p}_3)].\end{aligned}$$

The encoding channel Λ_σ is again purified to $\hat{U}_{\text{enc.}} = \exp(-i\sigma \hat{H}_{\text{enc.}})$ with $\hat{H}_{\text{enc.}} = \sqrt{2}\hat{x}_1 \hat{x}_2$ and mode 2 starting in vacuum. We also include two hiding unitaries: $\hat{U}_{\text{hide},1} = \exp(-i\sigma \hat{H}_{\text{hide},1})$ with $\hat{H}_{\text{hide},1} = g_1 \hat{x}_2 \hat{x}_3$, which removes the excess information due to the purification of the loss, and $\hat{U}_{\text{hide},2} = \exp(-i\sigma \hat{H}_{\text{hide},2})$ with $\hat{H}_{\text{hide},2} = \frac{1}{2} g_2 (\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2)$, which again removes the excess information due to the purification of the random displacement. Here, g_1 and g_2 are parameters to be minimized over again to obtain a tight bound.

The overall unitary is thus

$$\hat{U}_\sigma = \hat{U}_{\text{hide},2} \hat{U}_{\text{hide},1} \hat{U}_{\text{enc.}} \hat{U}_\eta,$$

the Hermitian generator of which is

$$\begin{aligned}\hat{H} = &\hat{U}_\eta^\dagger \hat{U}_{\text{enc.}}^\dagger \hat{U}_{\text{hide},1}^\dagger \hat{H}_{\text{hide},2} \hat{U}_{\text{hide},1} \hat{U}_{\text{enc.}} \hat{U}_\eta \\ &+ \hat{U}_\eta^\dagger \hat{U}_{\text{enc.}}^\dagger \hat{H}_{\text{hide},1} \hat{U}_{\text{enc.}} \hat{U}_\eta + \hat{U}_\eta^\dagger \hat{H}_{\text{enc.}} \hat{U}_\eta,\end{aligned}$$

where there is no term associated with $\hat{H}_\eta$, since $\hat{U}_\eta$ is independent of σ. This expression then simplifies to

$$\begin{aligned}\hat{H} = &\frac{g_2}{2}\left(\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2\right) \\ &+ (g_2\sigma + 1)\left(\sqrt{2\eta} + g_1\sqrt{1-\eta}\right)\hat{x}_3 \hat{x}_2 \\ &+ (g_2\sigma + 1)\left(\sqrt{2(1-\eta)} - g_1\sqrt{\eta}\right)\hat{x}_1 \hat{x}_2.\end{aligned}$$

The QFI of $\hat{U}_\sigma |\psi\rangle \otimes |\varphi\rangle$ is therefore

$$4\text{Var}[\hat{H}] = 2g_2^2 + f\ (g_1)\ (g_2\sigma + 1)^2$$

where, by Eq. (C5),

$$\begin{aligned}f\ (g_1) = &\left(\sqrt{2\eta} + g_1\sqrt{1-\eta}\right)^2 \\ &+ 2\langle \hat{x}_1^2 \rangle \left(\sqrt{2(1-\eta)} - g_1\sqrt{\eta}\right)^2.\end{aligned}$$

Minimizing over the free parameters of the hiding unitaries, g_1 and g_2, we find that

$$\begin{aligned}4\text{Var}[\hat{H}] &= \min_{g_1, g_2}\left[2g_2^2 + f\ (g_1)\ (g_2\sigma + 1)^2\right] \\ &= \min_{g_2}\left[2g_2^2 + \frac{4\langle \hat{x}_1^2 \rangle}{(1-\eta) + 2\eta\langle \hat{x}_1^2 \rangle}(g_2\sigma + 1)^2\right] \\ &= \frac{4}{2(\eta + \sigma^2) + (1-\eta)\langle \hat{x}_1^2 \rangle^{-1}}.\end{aligned}$$

Using the optimal value of $\langle \hat{x}_1^2 \rangle$ for a given constraint $\langle \hat{n} \rangle = N$ leads to the upper bound of Eq. (36). This optimal value of $\langle \hat{x}_1^2 \rangle$ for a fixed N would correspond to the state after the loss but before the encoding being an SMSV state, but there is no such state that becomes an SMSV state after a loss. Note that while, e.g., Schrödinger's cat states approach $\langle \hat{x}_1^2 \rangle = 2N$ for large N too, the relation $\langle \hat{x}_1^2 \rangle < \xi_N$ holds for any fixed N and any non-SMSV state. This bound, therefore, is not tight for a fixed finite N, as preparing an SMSV state does not saturate it. In the limit of $N \to \infty$, however, this bound proves the ECQFI in Eq. (35), which is saturated asymptotically by TMSV (see Appendix D) and thus is tight.

APPENDIX D: TWO-MODE SQUEEZED VACUUM

Here, we analyze the QFI from preparing a TMSV state given different noise models (loss and classical noise) and different encoding channels (one-dimensional and two-dimensional random-displacement channels).

In Ref. [37], it has been shown that the QFI of an M-mode Gaussian state with the parameter encoded in the

covariance matrix Σ is given by the following formula:

$$\mathcal{I}_Q(\sigma) = 2\mathrm{Tr}\left[(\partial_\sigma \Sigma)(4\Sigma \otimes \Sigma - \omega \otimes \omega)^{-1}(\partial_\sigma \Sigma)\right], \quad \text{(D1)}$$

where, if Σ is written in the quadrature basis of $(\hat{x}_1, \cdots, \hat{x}_M, \hat{p}_1, \cdots, \hat{p}_M)^{\mathrm{T}}$, then the commutator matrix is given by $\omega = \left[\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right]$, where $\mathbb{1}$ is the $M \times M$ identity matrix. This result generalizes Eq. (19) for single-mode Gaussian states.

For the one-dimensional random-displacement channel in Eq. (15) preceded by a loss channel, the QFI of a TMSV state with noisy ancilla $\eta_A > 0$ in the high-energy limit $N \to \infty$ and with $\sigma \ll 1$ is given by Eq. (D1) as

$$\mathcal{I}_Q(\sigma) = \frac{2(1-\eta_A)\sigma^2(\eta + \eta_A - 2\eta\eta_A)^2}{\xi + \sigma^2(\eta + \eta_A - 2\eta\eta_A)^3}, \quad \text{(D2)}$$

where

$$\xi = 2(1-\eta)\eta\eta_A\left(\eta^2(2(\eta_A - 1)\eta_A + 1) - 2\eta\eta_A^2 + \eta_A^2\right).$$

Attaining this QFI requires first antisqueezing and then, in general, applying a beam splitter followed by number-resolving measurements of the two modes [37]. For any $\sigma > 0$ and with noiseless ancilla ($\eta_A = 0$), the QFI in Eq. (D2) converges to the ECQFI in Eq. (35) in the high-energy limit. For example, in Fig. 4, we use the full form of the QFI for any σ, η, η_A, and N, which is too verbose to provide here but is found with the same method.

In comparison, for the two-dimensional random-displacement channel in Eq. (51) preceded by a loss channel, the high-energy QFI for a TMSV state with noisy ancilla is given by Eq. (D1) as

$$\mathcal{I}_Q(\sigma) = \frac{4(1-\eta_A)\sigma^2}{\left(\eta + \sigma^2\right)\left[(1-\eta)\eta_A + (1-\eta_A)\sigma^2\right]}. \quad \text{(D3)}$$

For $0 < \eta_A, \sigma^2 \ll \eta$, the two-dimensional encoding TMSV QFI in Eq. (D3) with ancilla loss η_A is approximately twice the one-dimensional encoding TMSV QFI in Eq. (D2) with ancilla loss $\eta_A/2$. In either case, the TMSV QFI attains the ECQFI at high energy if $\eta_A \ll \sigma^2, \eta$ but vanishes if $\sigma^2 \ll \eta, \eta_A$.

We now briefly analyze the perpendicular classical noise case discussed in Appendix J. Consider the one-dimensional random-displacement channel preceded by a perpendicular classical noise channel that adds σ_x to the system and $\sigma_{x,A}$ to the ancilla. Let us first consider the case of noiseless ancilla, i.e., $\sigma_{x,A} = 0$. Interestingly, for any $\langle \hat{n} \rangle = N$ per mode and any amount of perpendicular noise $\sigma_x > 0$, the QFI is

$$\mathcal{I}_Q(\sigma = 0) = \frac{8N(N+1)}{2N+1} \xrightarrow[N\to\infty]{} 4N, \quad \text{(D4)}$$

which does not exhibit the Rayleigh curse. Intuitively, for any N, there exists a symplectic transformation of $\hat{\rho}$, the state after the encoding channel, to a pair of modes such that one of them is completely decoupled from the perpendicular noise but nevertheless contains some information about the signal. An optimal measurement, therefore, would be to apply this symplectic transformation to $\hat{\rho}$ and then perform a number measurement of the perpendicular-noise–free mode. In the quadrature basis of $(\hat{x}_S, \hat{x}_A, \hat{p}_S, \hat{p}_A)^{\mathrm{T}}$, where the system S and ancilla A are entangled, the covariance matrix of $\hat{\rho}$ is

$$\Sigma = \frac{1}{2}\mathbb{1} + \begin{bmatrix} \sigma_x^2 + N & \sqrt{N(N+1)} \\ \sqrt{N(N+1)} & N \end{bmatrix} \oplus \begin{bmatrix} \sigma^2 + N & -\sqrt{N(N+1)} \\ -\sqrt{N(N+1)} & N \end{bmatrix}.$$

After performing a suitable antisqueezing operation, the inverse of the two-mode squeezing unitary that prepares the TMSV state from vacuum, the covariance matrix becomes

$$\Sigma = \frac{1}{2}\mathbb{1} + \frac{1}{2}\sigma_x^2\begin{bmatrix} N+1 & -\sqrt{N(N+1)} \\ -\sqrt{N(N+1)} & N \end{bmatrix} \oplus \frac{1}{2}\sigma^2\begin{bmatrix} N+1 & \sqrt{N(N+1)} \\ \sqrt{N(N+1)} & N \end{bmatrix}.$$

The collective mode with annihilation operator $\hat{a}_{\mathrm{dec}} = (1/\sqrt{2N+1})\left(\sqrt{N}\hat{a}_S + \sqrt{N+1}\hat{a}_A\right)$, therefore, is completely decoupled from the perpendicular noise (σ_x^2) yet remains coupled to the signal (σ^2). The covariance matrix of this perpendicular-noise–free mode is

$$\Sigma_{\mathrm{dec}} = \frac{1}{2}\mathbb{1} + \mathrm{diag}\left(0,\ 2\sigma^2\frac{N(N+1)}{2N+1}\right),$$

such that a number measurement of this mode attains the QFI in Eq. (D4). This protocol is reminiscent of the idea of displacement-noise–free interferometry in which there exist modes unaffected by displacement noise but that nevertheless contain information about the signal [140–142].

We now consider the case of noisy ancilla, i.e., $\sigma_{x,A} > 0$. Here, however, the QFI as $\sigma \to 0$ exhibits the Rayleigh curse by Claim 1 for a fixed value of N, since, for all i, $\bar{n}_i = c_i(\sigma_x, \sigma_{x,A}) + \mathcal{O}(\sigma^2)$ for some $c_i(\sigma_x, \sigma_{x,A}) > 0$ (see Appendix A). For a fixed finite signal $\sigma > 0$ and any fixed σ_x and $\sigma_{x,A}$, the QFI in Eq. (D1) attains the noiseless ECQFI of $2/\sigma^2$ in the high-energy limit of $N \to \infty$ particles per mode.

Finally, we elaborate on the relation between a TMSV state with noiseless ancilla and GKP states in the high-energy limit. The infinite-energy TMSV pure state is the

following unnormalized state:

$$\int_{-\infty}^{\infty} \mathrm{d}x\ |\hat{x}_S = x\rangle|\hat{x}_A = x\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\ |\hat{p}_S = p\rangle|\hat{p}_A = -p\rangle,$$

where $|\hat{x}_j = x_0\rangle$ ($|\hat{p}_j = p_0\rangle$) is the position (momentum) eigenstate of mode j at x_0 (p_0). We observe that deterministic or random displacements along $\hat{p}_S$ keep the state inside the code space of $\left\{|\hat{x}_S = x\rangle|\hat{x}_A = x\rangle\right\}_x$, where deterministic (random) displacement induces rotation (dephasing) inside this subspace. In comparison, deterministic or random displacements along $\hat{x}_S$ take the state outside of this code space and can be detected by measuring $\hat{x}_S - \hat{x}_A$. This is similar to the GKP infinite-energy state $|\mathrm{GKP}_{\mathrm{ideal}}\rangle$ from Appendix H, for which the code space is $\left\{\sum_{j=-\infty}^{\infty}|\hat{p} = 2j\sqrt{\pi} + p_0\rangle\right\}_{p_0}$. Deterministic (random) displacements along $\hat{p}$ induce rotation (dephasing) inside this subspace, while displacements along $\hat{x}$ take the state outside of this code space and can be detected by the relevant "syndrome" measurement of $\hat{x}$ mod $2\sqrt{\pi}$ [52].

APPENDIX E: PROOF OF CLAIM 4

Here, we prove Claim 4 by a Taylor expansion and again using the fact that, in the limit of small signals $\sigma \to 0$, an optimal measurement is to project onto the eigenbasis.

In general, a unitary $\hat{U}_\theta$ may be expanded around $\theta = 0$ to obtain the following approximation up to $\mathcal{O}(\theta^3)$:

$$\hat{U}_\theta \approx \hat{U}_0 + \theta\dot{\hat{U}}_0 + \frac{1}{2}\theta^2\ddot{\hat{U}}_0 \tag{E1}$$

where the unitarity condition $\hat{U}_\theta^\dagger\hat{U}_\theta = 1$ implies that

$$0 \approx \dot{\hat{U}}_0^\dagger\hat{U}_0 + \hat{U}_0^\dagger\dot{\hat{U}}_0 \tag{E2}$$

$$0 \approx \ddot{\hat{U}}_0^\dagger\hat{U}_0 + \hat{U}_0^\dagger\ddot{\hat{U}}_0 + 2\dot{\hat{U}}_0^\dagger\dot{\hat{U}}_0. \tag{E3}$$

Let $\hat{H} := i\hat{U}_0^\dagger\dot{\hat{U}}_0$ and $\hat{Z} := -\hat{U}_0^\dagger\ddot{\hat{U}}_0$ such that $\hat{Z} = \hat{H}^2 + 2i\hat{B}$, where $\hat{H}^2 = \mathrm{Re}[\hat{Z}] = \dot{\hat{U}}_0^\dagger\dot{\hat{U}}_0$ and $\hat{B} = \frac{1}{2}\mathrm{Im}[\hat{Z}] = (i/4)(\hat{U}_0^\dagger\ddot{\hat{U}}_0 - \ddot{\hat{U}}_0^\dagger\hat{U}_0)$.

Using this expansion around $\theta = 0$, the random unitary channel Λ_σ in Eq. (37) is approximately

$$\Lambda_\sigma(\hat{\rho}) \approx \hat{U}_0\hat{\rho}\hat{U}_0^\dagger + \sigma^2\left(\dot{\hat{U}}_0\hat{\rho}\dot{\hat{U}}_0^\dagger + \frac{1}{2}(\ddot{\hat{U}}_0\hat{\rho}\hat{U}_0^\dagger + \hat{U}_0\hat{\rho}\ddot{\hat{U}}_0^\dagger)\right).$$

Without loss of generality, we can study $\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0$, since $\hat{U}_0^\dagger$ does not depend on the parameter θ and thus will not affect the QFI if applied to the state after the channel. The final state then becomes

$$\begin{aligned}\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0 &\approx \hat{\rho} + \sigma^2\left(\hat{H}\hat{\rho}\hat{H} - \frac{1}{2}(\hat{Z}\hat{\rho} + \hat{\rho}\hat{Z}^\dagger)\right)\\ &= \hat{\rho} - i[\sigma^2\hat{B}, \hat{\rho}] + \sigma^2\left(\hat{H}\hat{\rho}\hat{H} - \frac{1}{2}\{\hat{H}^2, \hat{\rho}\}\right).\end{aligned} \tag{E4}$$

This is an approximation to the master-equation evolution of $\hat{\rho}$ by the Hamiltonian $\sigma^2\hat{B}$ and the Lindbladian jump operator $\hat{H}$ with decay rate σ^2 in the limit of $\sigma^2 T \ll 1$, where T is the total evolution time. In the additive case of $\hat{U}_\theta = \exp(-i\theta\hat{H})$, $\hat{B} = 0$ such that $\hat{Z} = \hat{H}^2$ and there is no Hamiltonian evolution. Moreover, Λ_σ is then a decoherence channel in the eigenbasis of $\hat{H}$ similarly to Eq. (16) for $\hat{H} = \hat{x}$.

Since $\partial_\sigma(\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0) \to 0$ as $\sigma \to 0$, the optimal measurement (after $\hat{U}_0^\dagger$) is projection $\hat{\Pi}$ onto the support of $\hat{\rho}$, i.e., $\hat{\Pi} = \sum_j|\psi_j\rangle\langle\psi_j|$ in the eigenbasis $\{|\psi_j\rangle\}_j$ of $\hat{\rho}$ [40,51]. In particular, since $\mathrm{Tr}[\hat{\Pi}[\sigma^2\hat{B}, \hat{\rho}]] = 0$, we may ignore the Hamiltonian evolution in the limit of $\sigma \to 0$. By Eq. (E4), the probability to remain in the support of $\hat{\rho}$ is

$$p = \mathrm{Tr}[\hat{\Pi}\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0] = 1 + \sigma^2\left(\langle\hat{H}\hat{\Pi}\hat{H}\rangle - \langle\hat{H}^2\rangle\right)$$

and the probability to leave is $1 - p = \sigma^2\langle\hat{H}\hat{\Pi}_\perp\hat{H}\rangle$, where $\hat{\Pi}_\perp = 1 - \hat{\Pi}$ is the projection onto the null space of $\hat{\rho}$. Here, the expectation values are calculated with respect to $\hat{\rho}$. Since the QFI equals the CFI of the optimal measurement,

$$\mathcal{I}_Q(\sigma = 0) = 4\langle\hat{H}\hat{\Pi}_\perp\hat{H}\rangle, \tag{E5}$$

and if $\hat{\rho}$ is pure, then the QFI is

$$\mathcal{I}_Q(\sigma = 0) = 4\mathrm{Var}[\hat{H}], \tag{E6}$$

which is equal to the QFI for the deterministic unitary case $\hat{U}_\sigma\hat{\rho}\hat{U}_\sigma^\dagger$. This correspondence between the deterministic and random cases does not hold for $\sigma > 0$. A similar observation for nonunitary channels in the limit of $\sigma \to 0$ has been made previously in Ref. [36].

APPENDIX F: ESTIMATING A WEAK DECAY RATE

Here, we consider the general scenario of estimating a weak decay rate. This is inspired by the fact that Eq. (E4) is analogous to a short-time solution of a master equation, where the jump operator is $\hat{H}$ and σ^2 is equivalent to the decay rate multiplied by the evolution time. We consider

the cases of a single jump operator and multiple jump operators sharing the same weak decay rate. This problem has previously been studied in Ref. [36], in the case of an initial pure state.

Suppose that a quantum state $\hat{\rho}(t)$ evolves in time t under the following Lindbladian master equation:

$$\partial_t\hat{\rho}(t)=\gamma\mathcal{L}_{\hat{Y}}[\hat{\rho}(t)],\quad \mathcal{L}_{\hat{Y}}[\hat{\rho}(t)]=\hat{Y}\hat{\rho}(t)\hat{Y}^\dagger-\frac{1}{2}\{\hat{Y}^\dagger\hat{Y},\hat{\rho}(t)\},$$

where $\gamma\geq 0$ is the decay rate, $\hat{Y}$ is the (potentially non-Hermitian) jump operator, and we seek to estimate $\sqrt{\gamma}$. In the short-time or weak-decay limit of $\gamma Tc\ll 1$, where T is the total evolution time and $c=\mathrm{Tr}[\hat{Y}\hat{\rho}\hat{Y}^\dagger]$. Then, the evolution of an initial state $\hat{\rho}$ is approximated by the channel Λ_γ, with the following Kraus representation:

$$\Lambda_\gamma(\hat{\rho})=\hat{K}_0\hat{\rho}\hat{K}_0^\dagger+\hat{K}_1\hat{\rho}\hat{K}_1^\dagger,$$

where the Kraus operators are

$$\hat{K}_0:=1-\frac{\gamma T}{2}\hat{Y}^\dagger\hat{Y}\approx\cos\left(\sqrt{\gamma T\hat{Y}^\dagger\hat{Y}}\right),$$
$$\hat{K}_1:=\sqrt{\gamma T}\hat{Y}\approx\sin\left(\sqrt{\gamma T}\hat{Y}\right),$$

such that

$$1=\hat{K}_0^\dagger\hat{K}_0+\hat{K}_1^\dagger\hat{K}_1+\mathcal{O}\left((\gamma T)^2\right).$$

Since the total evolution time T is known, we can measure the final state $\Lambda_\gamma(\hat{\rho})$ to estimate the weak decay rate $\sqrt{\gamma}$. Since $\partial_{\sqrt{\gamma}}\Lambda_\gamma\to 0$ as $\sqrt{\gamma}\to 0$, then the QFI with respect to $\sqrt{\gamma}$ is [40,51]

$$\mathcal{I}_Q(\sqrt{\gamma}=0)=4T\langle\hat{Y}^\dagger\hat{\Pi}_\perp\hat{Y}\rangle,\tag{F1}$$

where $\hat{\Pi}_\perp$ projects onto the null space of the state $\hat{\rho}=\sum_j p_j|\phi_j\rangle\langle\phi_j|$ before Λ_γ. If $\hat{\rho}=|\phi\rangle\langle\phi|$ is pure, then Eq. (F1) becomes

$$\mathcal{I}_Q(\sqrt{\gamma}=0)=4T\left(\langle\hat{Y}^\dagger\hat{Y}\rangle-|\langle\hat{Y}\rangle|^2\right),\tag{F2}$$

which resembles the formula for the variance $\mathrm{Var}[Z]=\mathrm{E}\left[|Z|^2\right]-|\mathrm{E}[Z]|^2$ of a complex random variable Z. If $\hat{Y}$ is Hermitian, then $\mathcal{I}_Q(\sqrt{\gamma}=0)=4T\mathrm{Var}[\hat{Y}]$, which is equivalent to Eq. (E5) for $\gamma T=\sigma^2$ and $\hat{Y}=\hat{H}$. If instead $\hat{\rho}$ is mixed, then Eq. (F1) implies that the Rayleigh curse will be avoided if and only if $\hat{\rho}$ is partial rank and the jump operator $\hat{Y}$ transitions some of $\hat{\rho}$ into its null space. $\hat{\rho}$ may be mixed, e.g., because Λ_γ is preceded by another noise channel $\Lambda_{\gamma'}$ with its own decay rate γ', which might not be small, and jump operator $\hat{Y}'$ such that the total channel is $\Lambda_\gamma\circ\Lambda_{\gamma'}$. It is perhaps more natural, however, to consider these two processes occurring simultaneously within the following Lindbladian master equation:

$$\partial_t\hat{\rho}(t)=\gamma\mathcal{L}_{\hat{Y}}[\hat{\rho}(t)]+\gamma'\mathcal{L}_{\hat{Y}'}[\hat{\rho}(t)].\tag{F3}$$

We consider an example of this scenario below.

1. Example: Qubit noise channels

Consider a qubit simultaneously undergoing amplitude damping (loss) and dephasing such that the Lindbladian master equation is thus

$$\partial_t\hat{\rho}(t)=\gamma_{\mathrm{deph}}\mathcal{L}\left(\hat{\sigma}_z\right)+\gamma_{\mathrm{loss}}\mathcal{L}\left(\hat{\sigma}_-\right).$$

After evolving for a time T, the individual solutions to the amplitude damping and dephasing channels are, respectively,

$$\Lambda_{\mathrm{deph}}\left(\hat{\rho}\right)=\begin{bmatrix}\rho_{1,1} & \rho_{1,0}e^{-\gamma_{\mathrm{deph}}T}\\ \rho_{1,0}^*e^{-\gamma_{\mathrm{deph}}T} & \rho_{0,0}\end{bmatrix},$$
$$\Lambda_{\mathrm{loss}}\left(\hat{\rho}\right)=\begin{bmatrix}\rho_{1,1}e^{-\gamma_{\mathrm{loss}}T} & \rho_{1,0}e^{-\gamma_{\mathrm{loss}}T/2}\\ \rho_{1,0}^*e^{-\gamma_{\mathrm{loss}}T/2} & \rho_{0,0}+\rho_{1,1}\left(1-e^{-\gamma_{\mathrm{loss}}T}\right)\end{bmatrix},$$

where $\rho_{j,k}=\langle j|\hat{\rho}|k\rangle$. Even though $[\hat{\sigma}_z,\hat{\sigma}_-]\neq 0$, these channels commute such that the solution to the Lindbladian master equation is given by $\Lambda_{\mathrm{deph}}\circ\Lambda_{\mathrm{loss}}\left(\hat{\rho}\right)=\Lambda_{\mathrm{loss}}\circ\Lambda_{\mathrm{deph}}\left(\hat{\rho}\right)$.

This means that estimating a weak loss $\sqrt{\gamma_{\mathrm{loss}}}\ll 1$ is unaffected by dephasing, since the optimal initial state for sensing Λ_{loss} is $|1\rangle\langle 1|$, which is stabilized by Λ_{deph}. By Eq. (F2), the QFI with respect to $\sqrt{\gamma_{\mathrm{loss}}}$ is thus $\mathcal{I}_Q(\sqrt{\gamma_{\mathrm{loss}}})=4T$.

On the other hand, estimating a weak dephasing $\sqrt{\gamma_{\mathrm{deph}}}\ll 1$ is strongly affected by loss. (We assume that $\gamma_{\mathrm{loss}}>0$ is fixed and known.) Eq. (F1) implies that the QFI for any initial state is Rayleigh cursed, since the state after Λ_{loss} is either full rank or $|0\rangle\langle 0|$, which has zero QFI. The QFI with respect to $\sqrt{\gamma_{\mathrm{deph}}}$ therefore vanishes in the limit of $\gamma_{\mathrm{deph}}\to 0$ for any single-qubit initial state.

2. Multiple jump operators

Suppose instead that the same weak decay rate γ is common to m jump operators $\{\hat{Y}_j\}_{j=1}^m$ such that the master equation is

$$\partial_t\hat{\rho}(t)=\gamma\sum_{j=1}^m\mathcal{L}_{\hat{Y}_j}[\hat{\rho}(t)].$$

Then, the evolution of $\hat{\rho}$ is approximated in the limit of $\gamma Tmc\ll 1$, where $c=\max_j\mathrm{Tr}[\hat{Y}_j\hat{\rho}\hat{Y}_j^\dagger]$, by the following

channel:

$$\Lambda_\gamma(\hat{\rho}) = \sum_{j=1}^{m} \left(\hat{K}_{0,j}\hat{\rho}\hat{K}_{0,j}^\dagger + \hat{K}_{1,j}\hat{\rho}\hat{K}_{1,j}^\dagger \right),$$

where the Kraus operators are

$$\hat{K}_{0,j} := \frac{1}{\sqrt{m}} - \frac{\gamma T\sqrt{m}}{2}\hat{Y}_j^\dagger \hat{Y}_j,$$
$$\hat{K}_{1,j} := \sqrt{\gamma T}\hat{Y}_j.$$

By a similar argument to the $m = 1$ case above, then

$$\mathcal{I}_Q(\sqrt{\gamma} = 0) = 4T\sum_{j=1}^{m}\langle \hat{Y}_j^\dagger \hat{\Pi}_\perp \hat{Y}_j\rangle. \tag{F4}$$

For example, in Eq. (51), $\gamma T = \sigma^2$, $\hat{Y}_1 = \hat{x}$, and $\hat{Y}_2 = \hat{p}$, such that Eq. (F4) implies Eq. (52).

APPENDIX G: FINITE-DIMENSIONAL SYSTEMS

Here, we consider a random unitary channel acting on a finite-dimensional system. Given a d-dimensional quantum system, the random unitary channel Λ_σ is defined analogously to Eq. (37) as $\Lambda_\sigma(\hat{\rho}) = \int_{-\infty}^{\infty} \mathrm{d}\theta\, p(\theta)\hat{U}_\theta\hat{\rho}\hat{U}_\theta^\dagger$, where $p(\theta) \sim \mathcal{N}(0,\sigma^2)$ and $\hat{U}_\theta$ is a θ-dependent unitary that acts on this system. We discuss the differences between finite-dimensional systems and the infinite-dimensional harmonic oscillator system that we study. In particular, we give an example where the analogous upper bound on the ECQFI to Eq. (32) can be loose for $\sigma > 0$.

For finite-dimensional systems, the optimal initial pure state in the limit of $\sigma \to 0$ is an equal superposition of the states corresponding to the smallest and largest eigenvalues of $\hat{H}$ to maximize $4\mathrm{Var}[\hat{H}]$. If a loss precedes the channel, then the mixed state $\hat{\rho}$ before the encoding needs to instead optimize $\langle \hat{H}\hat{\Pi}_\perp\hat{H}\rangle$.

For $\sigma > 0$, however, the same state may no longer be optimal, even in the lossless case. For example, consider sensing the phase diffusion of a qubit, modeled as a random rotation channel Λ_σ with $\hat{U}_\theta = \exp(-i\theta\hat{H})$ and $\hat{H} = \frac{1}{2}\hat{\sigma}_z$. The optimal initial state for $\sigma > 0$, even allowing for entangled resources, is, e.g., $|\uparrow_x\rangle$ with a QFI of $\sigma^2/(e^{\sigma^2} - 1)$, as this saturates the ECQFI as proved below. In comparison, the naive upper bound on the ECQFI, analogous to Claim 2, is $2/(\sigma^2 + 2)$, which is loose for $\sigma > 0$, unlike for the random bosonic displacement channel.

The ECQFI for the random rotation channel Λ_σ can be found as follows. The channel has the following Kraus operators:

$$\hat{K}_1 = \sqrt{\frac{1}{2}\left(1 + e^{-\frac{1}{2}\sigma^2}\right)}\hat{1}, \quad \hat{K}_2 = \sqrt{\frac{1}{2}\left(1 - e^{-\frac{1}{2}\sigma^2}\right)}\hat{\sigma}_z,$$

such that an upper bound on the ECQFI is [139]

$$\mathcal{I}_Q^{\Lambda_\sigma}(\sigma) \le 4\left\| \sum_j \dot{\hat{K}}_j^\dagger \dot{\hat{K}}_j \right\| = \frac{\sigma^2}{e^{\sigma^2} - 1}.$$

This implies that $|\uparrow_x\rangle$ is optimal and attains the ECQFI.

APPENDIX H: REVIEW OF FINITE-ENERGY GKP STATES

Here, we review the families of finite-energy states that we study, which approximate the ideal GKP state in the high-energy limit.

The ideal GKP grid state is an unnormalized infinite-energy single-mode state consisting of an equal superposition of infinitely many evenly spaced position eigenstates, e.g.,

$$|\mathrm{GKP}_{\mathrm{ideal}}\rangle \propto \sum_{j=-\infty}^{\infty} |\hat{x} = 2j\sqrt{\pi}\rangle = \sum_{j=-\infty}^{\infty} |\hat{p} = 2j\sqrt{\pi}\rangle,$$

where $|\hat{x} = x_0\rangle$ ($|\hat{p} = p_0\rangle$) is the position (momentum) eigenstate at x_0 (p_0) [52]. There are many different continuous families of normalized finite-energy states that approximate the same ideal state in the high-energy limit [143]. One such family, $|\mathrm{GKP}_\Delta\rangle$, is defined as a superposition with a Gaussian window function of $\mathcal{N}(0, \frac{1}{2}\Delta^{-2})$ displaced squeezed single-mode states, with variance $\frac{1}{2}\Delta^2$ and evenly spaced by $2\sqrt{\pi}$, as follows:

$$|\mathrm{GKP}_\Delta\rangle \propto \sum_{j=-\infty}^{\infty} \exp\left(-2\pi\Delta^2 j^2\right)\hat{U}_{2j\sqrt{\pi}}|\mathrm{SMSV}_{\frac{1}{2}\Delta}\rangle,$$

where $\hat{U}_\mu = \exp(-i\mu\hat{p})$ translates $\hat{x}$ to $\hat{x} + \mu$ and the position-basis wave function of the SMSV state is

$$|\mathrm{SMSV}_{\frac{1}{2}\Delta}\rangle = \frac{1}{\pi^{\frac{1}{4}}\sqrt{\Delta}}\int_{-\infty}^{\infty}\mathrm{d}x\ \exp\left(-\frac{x^2}{2\Delta^2}\right)|\hat{x} = x\rangle.$$

This family approximates the ideal GKP state, $|\mathrm{GKP}_{\mathrm{ideal}}\rangle$, in the limit of $\Delta \to 0$. Numerically, for a fixed but small $\Delta > 0$, $|\mathrm{GKP}_\Delta\rangle$ can be approximated by a finite sum over $j = -J, \cdots, J$, where some large integer J is chosen such that $2\pi\Delta^2 J^2 \gg 1$.

The Wigner function of the pure initial state $|\mathrm{GKP}_\Delta\rangle$ is shown in Fig. 7(a), where the marginals fit the Gaussian envelope of $\mathcal{N}(0, \frac{1}{2}\Delta^{-2})$. The mixed state after a loss,

$\hat{\rho} = \Lambda_\eta^{\text{loss}}(|\text{GKP}_\Delta\rangle\langle\text{GKP}_\Delta|)$, is shown in Fig. 7(b), where, visually, the subvacuum peaks grow and move toward the origin. Solely for the purpose of plotting, we calculate these Wigner functions using a different family of finite-energy GKP states $|\text{GKP}'_\varepsilon\rangle$, defined as the normalized result of applying the nonunitary operator $\exp(-\varepsilon\hat{n})$ to the ideal GKP state [143]. This approximation is valid since, when $\frac{1}{24}\Delta^3$ is negligible compared to $\frac{1}{2}\Delta$, then $|\text{GKP}_\Delta\rangle \approx |\text{GKP}'_{\varepsilon=\Delta^2}\rangle$, e.g., this holds for the value of $\varepsilon = 0.01$ used in Fig. 7. We use this family $|\text{GKP}'_\varepsilon\rangle$ here since the Wigner function is more efficient to calculate than in the Fock-basis representation shown in Figs. 7(c)–7(d) [144]. While, at low energies, the different families of finite-energy GKP states diverge, e.g., the behavior of $|\text{GKP}_\Delta\rangle$ as $N \to 0$ seen in Fig. 4(b) may be different from $|\text{GKP}'_{\varepsilon=\Delta^2}\rangle$, we are principally interested in the high-energy limit where they agree.

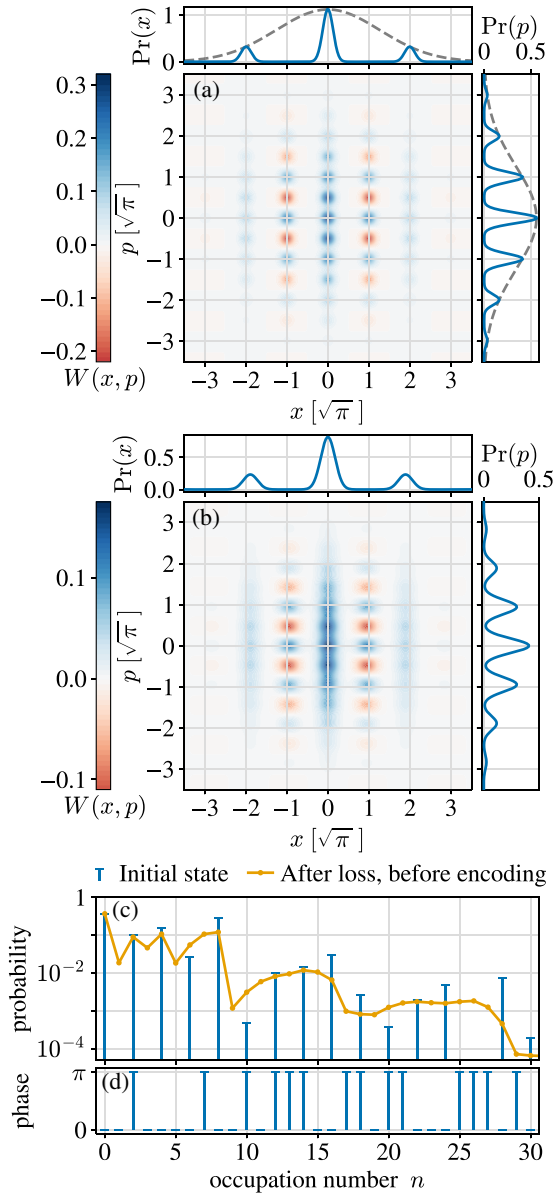

FIG. 7. The GKP finite-energy state for $\Delta^2 \approx 0.1$: Wigner-function and quadrature probability distributions of (a) the pure initial state and (b) the state after a loss of $\eta = 0.1$ before the encoding. The Gaussian envelope of the marginals for the pure initial state is shown. Visually, the loss channel enlarges each subvacuum peak and pulls it toward the origin. (c) Fock-basis probabilities on a logarithmic scale shown up to $n = 30$ before and after the loss, in a truncated Hilbert space of dimension 100. (d) Fock-basis complex phases of the pure initial state. The coefficients are all real.

APPENDIX I: NUMERICAL METHODS

Here, we discuss the numerical methods mentioned in Sec. V F that we use to search for the optimal initial single-mode state in the presence of loss. We first discuss our brute-force approach and then an alternative approach that we have experimented with.

1. Brute-force approach

While we have observed numerically that preparing finite-energy GKP states approaches the ECQFI for $0 < \sigma^2 \ll \eta, \eta_A$ in the limit of large $\langle\hat{n}\rangle = N$, we do not know if this is the only such family of non-Gaussian states. Moreover, while we know that the upper bound in Eq. (36) is loose for a fixed finite $\langle\hat{n}\rangle = N$, we do not know whether a TMSV state with noiseless ancilla is optimal. We search here numerically for non-Gaussian states that might perform better.

We consider sparse superpositions of finitely many Fock states. For simplicity, we consider superpositions equally spaced in the number basis, i.e., $|\psi\rangle = \sum_{k=0}^{K-1} c_k|mk\rangle$ for some spacing m and number of peaks K. For successively larger truncations of the Hilbert space in the number basis, we fix different values of m and K and numerically optimize the basis coefficients $\{c_k\}_{k=0}^{K-1}$ to maximize the QFI using gradient-descent and particle-swarm optimization methods. This is a brute-force approach to finding the optimal state.

We show an example of one such state in Fig 8, which has been found numerically within a truncated Hilbert space of dimension 490. This state, $|\psi_{\text{num.}}\rangle = \sum_{j=0}^{23} c_j|20j\rangle$, where $m = 20$ and $K = 24$, has a QFI of 18.4 for $\langle\hat{n}\rangle = 158.9$, which is within 9% of the ECQFI of 20 for $\sigma^2 = 10^{-6}$ and $\eta = 0.1$. (This value of the QFI appears stable under small perturbations in σ and η.) It is important here that the signal and loss levels are fixed and finite, since the intuition for this sparse state is that, e.g., the transition to $|1\rangle$ is dominated by the signal trajectory from $|0\rangle$ rather than the loss trajectory from $|20\rangle$, as shown in Fig. 8(c). If the loss was larger or the signal was smaller, then we would expect a larger separation m to be necessary. The optimal measurement for this state

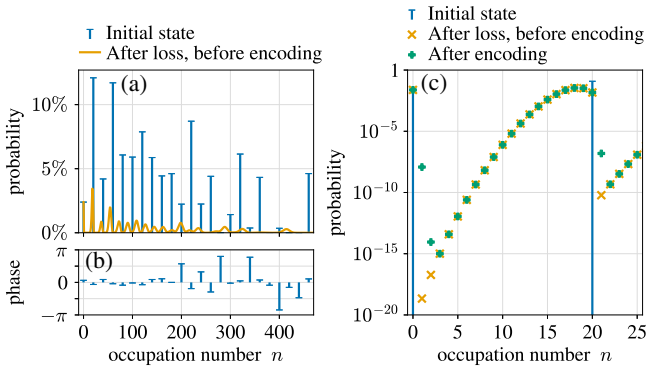

FIG. 8. The Fock-basis (a) probabilities and (b) complex phases of the non-Gaussian initial state $|\psi_{\text{num.}}\rangle$ found by numerical optimization. The probabilities after the loss of $\eta = 0.1$ but before the encoding are a sum of binomial distributions. (c) An example of two neighboring components. For $\sigma = 10^{-3}$, the transition to $|1\rangle$ is dominated by the signal trajectory $|0\rangle \mapsto |1\rangle$ rather than the loss trajectory $|20\rangle \mapsto |1\rangle$. However, this is not the case for all neighboring components of $|\psi_{\text{num.}}\rangle$.

is a superposition of number-resolving measurements, i.e., projections onto different linear combinations of the Fock states, similar to the generalized parity measurements for binomial codes [60]. We conjecture that optimizing over superpositions of unequally spaced but still sparsely separated Fock states in larger truncations of the Hilbert space can get arbitrarily close to the ECQFI for any $\sigma^2 \ll \eta$.

2. Biconvex optimization

Here, we describe this task as a biconvex-optimization problem and discuss how, in principle, it might be solved without using a brute-force approach.

The general problem of finding the optimal protocol, i.e., the initial state and measurement scheme, for sensing a parameter θ encoded by a quantum channel Λ_θ may be formulated as follows. This problem reduces to biconvex optimization of the function $f(\hat{\rho}, X)$, given by [145,146]

$$f(\hat{\rho}, X) := -\text{Tr}\Big[\hat{\rho}\Big(-\Lambda_\theta^\dagger(X^2) + \dot{\Lambda}_\theta^\dagger(X)\Big)\Big],$$

where $\Lambda_\theta^\dagger$ is the conjugate channel of Λ_θ in the Heisenberg-picture sense, $\dot{\Lambda}_\theta^\dagger = \mathrm{d}\Lambda_\theta^\dagger/\mathrm{d}\theta$, and $X \in \mathcal{L}^2(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ with finite dimension (e.g., on a truncated bosonic Hilbert space). Here, the biconvexity of f means that f is convex with respect to $\hat{\rho}$ if X is held constant and with respect to X if $\hat{\rho}$ is held constant, but not necessarily with respect to both at once. The global minimum of f exists and is the negative of the CQFI in Eq. (28),

$$\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta) = -\inf_X \inf_{\hat{\rho}} f(\hat{\rho}, X), \tag{I1}$$

where the order of the infima may be exchanged and the negative sign is included as a convention such that f is biconvex rather than biconcave. If X is the SLD of $\Lambda_\theta(\hat{\rho})$ with respect to θ, then $f(\hat{\rho}, X) = -\mathcal{I}_Q^{\Lambda_\theta(\hat{\rho})}$ is the negative of the QFI given the initial state $\hat{\rho}$. This biconvex formulation can also be applied to optimize the CQFI subject to a convex constraint. For example, the average-energy–constrained CQFI is given by

$$\mathcal{I}_Q^{\Lambda_\theta,N,\text{no ancilla}}(\theta) = -\inf_X \inf_{\hat{\rho}} f(\hat{\rho}, X) \quad \text{subject to } \text{Tr}\big[\hat{n}\hat{\rho}\big] = N. \tag{I2}$$

This is a biconvex-optimization problem since the set of density matrices that satisfy this average-energy constraint of $\text{Tr}\big[\hat{n}\hat{\rho}\big] = N$ is convex.

In Ref. [145], these biconvex-optimization problems have been proposed to be solved locally using an alternating convex search (ACS) [147]. The ACS proceeds from a given starting point $\hat{\rho}_0$ by finding the corresponding SLD X_0 that solves the convex problem with $|\psi_0\rangle$ held constant, then finds the corresponding $\hat{\rho}_1$ that solves the convex problem with X_0 held constant, and so on. Although the ACS is monotonic and converges, it is only a local search and is not guaranteed to find the global minimum, although different random starting points $\hat{\rho}_0$ can be used to explore the nonconvex geometry [147].

In theory, a deterministic global optimization program (GOP) exists for solving biconvex problems [148]. The GOP, however, only converges in exponential time in the number of confounding variables. Here, this would be proportional to the dimension of the truncated Hilbert space squared in the worst case. Although methods exist to reduce the number of confounding variables in particular cases [149], the GOP is impractical in general. This limitation of the GOP and success of the ACS has previously been discussed in the context of other biconvex-optimization problems in quantum information science (see, e.g., Refs. [136,146,150–152]).

While the ACS has proved successful in these other applications, e.g., it typically converges very fast to optimal values from arbitrary initial states, it appears to be less efficient for our problem of sensing σ from Λ_σ. We have implemented this method for calculating both the unconstrained CQFI in Eq. (I1) and the average-energy–constrained CQFI in Eq. (I2). In the unconstrained CQFI case, we have observed a strong dependence on the starting point. Starting from the vacuum state, the ACS does not converge to or surpass the QFI of the finite-energy GKP states even after 10 000 iterations. This method also does not significantly improve upon the finite superpositions of Fock states from the brute-force approach when started from them. We observe improved performance when running the ACS with random pure states. (We sample the real parts of the number-basis coefficients from a

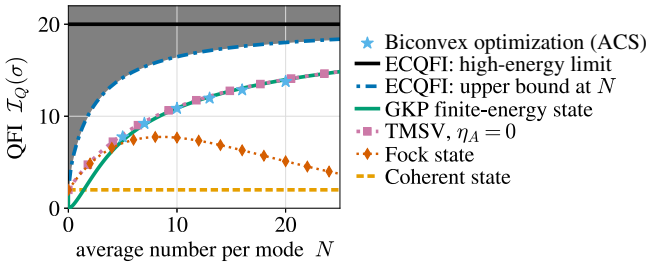

FIG. 9. The QFI versus the initial average occupation number per mode for the states obtained from biconvex optimization via the ACS with the average energy-constraint of Eq. (I2) after 500 iterations and starting from random pure states. The remaining curves corresponding to the QFI from different states are the same as the lossy case in Fig. 4(b).

Gaussian distribution with zero mean and unit variance and then normalize the state. The imaginary parts are zero.) However, the ACS still fails to converge to the QFI of the TMSV and GKP states at finite average energy. Determining what aspects of the geometry of our Rayleigh-cursed bosonic problem limit the ACS is left to future work.

Implementing the ACS for the average energy–constrained CQFI case in Eq. (I2) leads to considerably improved results: the ACS reaches values that are $\geq 98\%$ of the QFI of a finite-energy TMSV after 500 iterations for different N. These results are shown in Fig. 9. This provides strong numerical evidence for our conjecture that the finite-energy CQFI corresponds to the finite-energy TMSV QFI. These results also suggest that the failure of the ACS in the unconstrained CQFI case may be due to the bosonic nature of this problem: we observe that the unconstrained ACS on the truncated Hilbert space does not converge to the QFI of the finite-energy TMSV or GKP states, but requires an average-energy constraint that is much smaller than the truncation does. The states that the ACS finds are not obviously related to the TMSV or GKP states. We defer understanding the full space of states with similar QFI at the same average energy to future work.

APPENDIX J: CLASSICAL NOISE AND THE OPTIMAL INITIAL STATE

Here, we prove the results discussed in Sec. V G about the optimal initial state in the presence of significant classical noise but no loss.

Suppose that the encoding channel is preceded by the classical noise channel $\Lambda^{\text{noise}}_{\Sigma_C}$ in Eq. (22) for a given classical noise covariance matrix Σ_C such that the total channel is $\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda^{\text{noise}}_{\Sigma_C}$. We consider the following three cases in turn: $\Sigma_C = \text{diag}\left(0, \sigma_p^2\right)$, where the random displacements from the classical noise are parallel to (i.e., in the same quadrature, $\hat{p}$, as) those from the signal; $\Sigma_C = \text{diag}\left(\sigma_x^2, 0\right)$, where the classical noise is perpendicular (i.e., in the opposite quadrature, $\hat{x}$) to the signal; and $\Sigma_C = \text{diag}\left(\sigma_x^2, \sigma_p^2\right)$, where the classical noise is in both quadratures, which includes the isotropic case $\Sigma_C = \text{diag}\left(\sigma_C^2, \sigma_C^2\right)$. Whether the Rayleigh curse arises depends on Σ_C as shown below.

1. Parallel classical noise

For $\Sigma_C = \text{diag}\left(0, \sigma_p^2\right)$, we prove that the Rayleigh curse is unavoidable and that the optimal initial state for a given $\langle \hat{n} \rangle = N$ is to prepare an SMSV state.

We first note that the total channel corresponds to a combined classical noise channel:

$$\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma_p^2\right)} = \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma^2+\sigma_p^2\right)}, \tag{J1}$$

where $\Sigma_C = \text{diag}(0, \sigma^2 + \sigma_p^2)$ adds the signal and noise in quadrature since they are uncorrelated.

Equation (J1) holds immediately for Gaussian states, but let us prove it for any initial state $\hat{\rho}$ by direct calculation:

$$\begin{aligned}\Lambda'_\sigma(\hat{\rho}) &= \int_{\mathbb{R}^2} \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \, p_{\sigma^2}(\alpha_2) \, p_{\sigma_p^2}(\alpha_1) \, \hat{U}_{\alpha_2} \hat{U}_{\alpha_1} \hat{\rho} \hat{U}^\dagger_{\alpha_1} \hat{U}^\dagger_{\alpha_2} \\ &= \int_{\mathbb{R}^2} \mathrm{d}\alpha_1 \mathrm{d}\beta \, p_{\sigma^2}(\beta - \alpha_1) \, p_{\sigma_p^2}(\alpha_1) \, \hat{U}_\beta \hat{\rho} \hat{U}^\dagger_\beta,\end{aligned}$$

where $p_{\varsigma^2}(\alpha) \sim \mathcal{N}(0, \varsigma^2), \beta = \alpha_1 + \alpha_2$, and $\hat{U}_\alpha = \exp(i\alpha\hat{x})$. By recognizing the convolution between two Gaussian distributions, then this equals

$$\begin{aligned}&= \int_{\mathbb{R}} \mathrm{d}\beta \, p_{\sigma^2+\sigma_p^2}(\beta) \, \hat{U}_\beta \hat{\rho} \hat{U}^\dagger_\beta \\ &= \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma^2+\sigma_p^2\right)}(\hat{\rho}).\end{aligned}$$

By the chain rule, Eq. (J1) implies that the QFI for a given initial state $\hat{\rho}$ is

$$\mathcal{I}_Q^{\Lambda'_\sigma(\hat{\rho})}(\sigma) = \frac{\sigma^2}{\sigma^2 + \sigma_p^2} \mathcal{I}_Q^{\Lambda_\sigma(\hat{\rho})}\left(\sigma = \sqrt{\sigma^2 + \sigma_p^2}\right). \tag{J2}$$

Hence, the ECQFI $\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma)$ for a given value of σ_p is proportional to the noiseless ECQFI $\mathcal{I}_Q^{\Lambda_\sigma}(\sigma)$ evaluated at the signal with variance $\sigma^2 + \sigma_p^2$, which implies Eq. (38). Since preparing an SMSV state is the optimal initial state that attains the noiseless ECQFI for a fixed $\langle \hat{n} \rangle = N$, it is also the optimal initial state in the presence of parallel classical noise. The conversion factor of $\sigma^2/(\sigma^2 + \sigma_p^2)$, however, will introduce the Rayleigh curse for $\sigma \ll \sigma_p$ and mean that the noiseless ECQFI cannot be recovered regardless of the initial state.

Before moving on to discuss the other possible cases of classical noise, we analyze what happens when both parallel classical noise and loss are present. Let the total channel be

$$\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma_p^2\right)} \circ \Lambda^{\text{loss}}_\eta.$$

By combining Eqs. 35 and (J2), the ECQFI is then given in Eq. (39). By a similar argument to that made below Eq. (J2), the optimal initial states are the same as for the case with loss but no classical noise.

2. Perpendicular classical noise

For $\Sigma_C = \text{diag}(\sigma_x, 0)$, we prove that, for a fixed finite signal $\sigma > 0$, the noiseless ECQFI can be recovered in the high-energy limit and there exist finite-energy states that do not exhibit the Rayleigh curse as $\sigma \to 0$.

Suppose that we prepare an SMSV state with $\langle \hat{n} \rangle = N$, then, given perpendicular classical noise, the QFI calculated from Eq. (19) is

$$\mathcal{I}_Q(\sigma) = \frac{4}{(2\xi_N)^{-2}\left(\frac{\sigma_x}{\sigma}\right)^2 \frac{\left(\frac{1}{2}\sigma_x^2+\xi_N\right)}{\left(\sigma_x^2+\xi_N\right)^2} + \xi_N^{-1} + 2\sigma^2}$$
$$\xrightarrow[N\to\infty]{} \frac{4}{\xi_N^{-1} + 2\sigma^2}.$$

For a fixed finite signal $\sigma > 0$, the above limit as $N \to \infty$ implies that the noiseless ECQFI in Eq. (32) is recovered for large enough N such that $\sigma\xi_N \gg \sigma_x$. Note that the optimal measurement for an SMSV state in the noiseless case has been to first antisqueeze and then perform a number-resolving measurement. Here, intuitively, this antisqueezing operation squeezes the perpendicular noise such that its effect becomes vanishing in the high-energy limit.

Note that we assume a particular order of limits above: $\lim_{\sigma\to 0} \sup_{|\Psi\rangle} \mathcal{I}_Q(\sigma)$. As discussed in Sec. V E, this is the relevant order in practice because we are interested in small but finite signals $\sigma > 0$. For completeness, we now briefly discuss the opposite order of limits: $\sup_{|\Psi\rangle} \lim_{\sigma\to 0} \mathcal{I}_Q(\sigma)$. While the family of SMSV states is optimal as $N \to \infty$ for a fixed finite $\sigma > 0$, a given SMSV state with fixed $\langle \hat{n} \rangle = N$ exhibits the Rayleigh curse as $\sigma \to 0$ by Claim 1.

Unlike the parallel classical noise case, however, the Rayleigh curse can be avoided using different initial states. For example, we show in Appendix D that preparing a TMSV state with noiseless ancilla, i.e., no noise channels on the ancilla, with a fixed $\langle \hat{n} \rangle = N$ particles per mode yields a QFI of

$$\mathcal{I}_Q(\sigma = 0) = \frac{8N(N+1)}{2N+1} \xrightarrow[N\to\infty]{} 4N, \tag{J3}$$

which is independent of the amount of perpendicular classical noise. This recovers the noiseless QFI of $4N$ for a TMSV state in the high-energy limit, but not the noiseless ECQFI of $4\xi_N \approx 8N$ at $\sigma = 0$. Numerical analysis suggests that finite-energy GKP states are also resilient to perpendicular classical noise.

We note that the above analysis, while promising, does not include the relevant real-world case of perpendicular noise jointly in the presence of optical loss. There, the degradation of the squeezed state from the loss may prevent the complete removal of perpendicular noise.

3. Classical noise in both quadratures

For $\Sigma_C = \text{diag}\left(\sigma_x^2, \sigma_p^2\right)$ with $\sigma_x, \sigma_p > 0$, we combine the above results to show that the Rayleigh curse is inevitable, the noiseless ECQFI cannot be recovered, and that an SMSV state remains the optimal initial state for a fixed $\sigma > 0$ in the high-energy limit. These results apply, e.g., to the isotropic case of $\Sigma_C = \text{diag}\left(\sigma_C^2, \sigma_C^2\right)$ with $\sigma_C > 0$.

The random-displacement channels $\Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)}$ and $\Lambda^{\text{noise}}_{\text{diag}(0,\sigma_p^2)}$ commute with each other and obey the following relation:

$$\Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,\sigma_p^2)} = \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)} \circ \Lambda^{\text{noise}}_{\text{diag}(0,\sigma_p^2)}$$
$$= \Lambda^{\text{noise}}_{\text{diag}(0,\sigma_p^2)} \circ \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)},$$

such that, by Eq. (J1), the total channel is

$$\Lambda'_\sigma = \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,\sigma^2+\sigma_p^2)}.$$

Hence, similarly to Eq. (J2), the QFI for a given initial state $\hat{\rho}$ is

$$\mathcal{I}_Q^{\Lambda'_\sigma(\hat{\rho})}(\sigma) = \frac{\sigma^2}{\sigma^2+\sigma_p^2} \mathcal{I}_Q^{\left(\Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)}\right)(\hat{\rho})} \left(\sigma = \sqrt{\sigma^2+\sigma_p^2}\right).$$

The optimal initial state here, therefore, is the same as for the perpendicular case in which the total channel is $\Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)}$. For a fixed finite $\sigma > 0$, this implies that preparing an SMSV state is optimal in the high-energy limit but does not recover the noiseless ECQFI due to the $\sigma^2/(\sigma^2+\sigma_p^2)$ factor. In the limit of $\sigma \ll \sigma_p$, the same factor leads to an unavoidable Rayleigh curse.

When we restrict ourselves to study signals above a classical noise floor, e.g., in Sec. V E, we mean signals σ above an isotropic classical noise σ_C, i.e., for $\sigma \gg \sigma_C$, since all protocols exhibit the Rayleigh curse below the noise floor, i.e., for $\sigma \ll \sigma_C$.

APPENDIX K: ADAPTIVE PROTOCOL

Here, we detail an adaptive protocol for our separable measurements scheme (quadrature and number measurements) that converges to a factor of 2 away from the QFIM bound, as shown in Fig. 6(b).

Algorithmically, the adaptive measurement protocol proceeds as follows:

$k \leftarrow 1$
while $k \leq M$ **do**
  $p_k \leftarrow$ outcome of quadrature measurement $\hat{p}$
  Displace quadrature $\hat{p}$ by $-\frac{p_k}{k}$
  $n_k \leftarrow$ outcome of number measurement $\hat{n}$
  $k \leftarrow k+1$
end while

Let us denote $\langle \hat{p} \rangle$ after the kth step as μ_k. Since $p_k = \mu_{k-1} + n_k$, where n_k is a Gaussian random variable $n_k \sim \mathcal{N}\left(0, \frac{1}{2}\right)$, then, by induction, $\mu_k = -(1/k)\sum_{i=1}^{k} n_i$. This implies that $\Delta\mu_k = 1/\sqrt{2k}$, i.e., the "size" of μ_k is equal to the statistical uncertainty in estimating μ after k.

After all of the measurements are performed, we then estimate μ_i with $\tilde{\mu}_i = \mu_i - \mu_M$ and estimate σ with $\varsigma = \sqrt{(2/M)\sum_i n_i - (1/M)\sum_i \tilde{\mu}_i^2}$.

APPENDIX L: PROOF OF CLAIM 5

Here, we prove Claim 5. We first calculate the QFI and then explicitly construct the joint measurement, which is asymptotically optimal.

In the limit of $\sigma \to 0$, the QFI with respect to σ for a pure initial state is $4\mathrm{Var}[\hat{H}]$ by Claim 4. We can assume, without loss of generality, that $\mu \ll 1$. Then, the QFI with respect to μ can be shown to also be $4\mathrm{Var}[\hat{H}]$ by expanding the channel $\Lambda_{\sigma,\mu}$ in Eq. (46) to second order in μ and σ, since higher orders will not contribute to the QFI.

We now construct a joint measurement the CFIM of which saturates these QFIs simultaneously. Without loss of generality, we assume that $\langle \hat{H} \rangle = 0$ and let $V := \mathrm{Var}[\hat{H}]$ for the initial pure state $|\psi\rangle$. Then, $|\phi_1\rangle := (1/\sqrt{V})\hat{H}|\psi\rangle$ and $|\phi_2\rangle := (1/\sqrt{l})(\hat{H}^2 - V)|\psi\rangle$ are orthogonal to $|\psi\rangle$, where l is a normalization factor.

The final state $\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|)$ can be expanded up to second order in μ and σ using this orthonormal set of states $\{|\psi\rangle, |\phi_1\rangle, |\phi_2\rangle\}$ as

$$\begin{pmatrix} 1-\left(\mu^2+\sigma^2\right)V & -i\mu\sqrt{V} & -\frac{1}{2}\left(\mu^2+\sigma^2\right)l \\ i\mu\sqrt{V} & \left(\mu^2+\sigma^2\right)V & 0 \\ -\frac{1}{2}\left(\mu^2+\sigma^2\right)l & 0 & 0 \end{pmatrix},$$

such that we can restrict attention to the two-dimensional subspace of $|\psi\rangle$ and $|\phi_1\rangle$.

We want to consider joint measurements on M copies of the final state $\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|)$. Consider the following collective states:

$$|e_i\rangle := |\psi\rangle^{\otimes(i-1)}|\phi_1\rangle|\psi\rangle^{\otimes(M-i)}.$$

Let the symmetric state be $|e_s\rangle := (1/\sqrt{M})\sum_i |e_i\rangle$ and let the antisymmetric states be denoted as $\{|e_{a,i}\rangle\}_{i=1}^{M-1}$. In this collective basis, the final state is

$$\begin{aligned}\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|)^{\otimes M} \approx{}& \left[1 - MV(\mu^2+\sigma^2)\right]|\psi\rangle^{\otimes M}\langle\psi|^{\otimes M} \\ &- i\mu\sqrt{MV}\left(|e_s\rangle\langle\psi|^{\otimes M} - |\psi\rangle^{\otimes M}\langle e_s|\right) \\ &+ V\left(\sigma^2+\mu^2 M\right)|e_s\rangle\langle e_s| \\ &+ \sigma^2 V\sum_{i=1}^{M-1}|e_{a,i}\rangle\langle e_{a,i}|.\end{aligned}$$

We now calculate the CFIM obtained with projective measurement onto $(1/\sqrt{2})\left(|\psi\rangle^{\otimes M} + i|e_s\rangle\right)$ and $\left\{|e_{a,i}\rangle\right\}_{i=1}^{M-1}$ and show that it is asymptotically optimal. The probability to be projected onto the antisymmetric subspace is $\sigma^2 V(M-1)$, which provides the information about σ, while the probability to be projected onto $(1/\sqrt{2})\left(|\psi\rangle^{\otimes M} + i|e_s\rangle\right)$ is

$$p = \frac{1}{2}\left(1 - \sigma^2(M-1)V\right) + \sqrt{MV}\mu,$$

which provides the information about μ but no information about σ in the limit of $\sigma \to 0$. The CFIM with respect to μ and σ in the limit of $\sigma \to 0$ is

$$\mathcal{I}_C = \begin{pmatrix} 4MV & 0 \\ 0 & 4(M-1)V \end{pmatrix},$$

which saturates the QFIM and implies that this joint protocol is asymptotically optimal.

[1] E. P. Verlinde and K. M. Zurek, Observational signatures of quantum gravity in interferometers, Phys. Lett. B **822**, 136663 (2021).

[2] D. Li, V. S. H. Lee, Y. Chen, and K. M. Zurek, Interferometer response to geontropic fluctuations, Phys. Rev. D **107**, 024002 (2023).

[3] L. McCuller, Single-photon signal sideband detection for high-power Michelson interferometers, arxiv:2211.04016 [hep-ex, physics:physics, physics:quant-ph].

[4] A. Chou, H. Glass, H. R. Gustafson, C. Hogan, B. L. Kamai, O. Kwon, R. Lanza, L. McCuller, S. S. Meyer, J. Richardson, C. Stoughton, R. Tomlin, and R. Weiss, The holometer: An instrument to probe Planckian quantum geometry, Class. Quantum Grav. **34**, 065005 (2017).

[5] S. M. Vermeulen, L. Aiello, A. Ejlli, W. L. Griffiths, A. L. James, K. L. Dooley, and H. Grote, An experiment for

observing quantum gravity phenomena using twin tabletop 3D interferometers, Class. Quantum Grav. **38**, 085008 (2021).

[6] J. D. Romano and Neil J. Cornish, Detection methods for stochastic gravitational-wave backgrounds: A unified treatment, Living Rev. Relativ. **20**, 2 (2017).

[7] B. P. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. Adya, C. Affeldt, M. Agathos *et al.*, Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA, Living Rev. Relativ. **23**, 1 (2020).

[8] J. Aasi *et al.*, Advanced LIGO, Class. Quantum Grav. **32**, 074001 (2015).

[9] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin *et al.*, Advanced Virgo: A second-generation interferometric gravitational wave detector, Class. Quantum Grav. **32**, 024001 (2015).

[10] T. Akutsu, M. Ando, K. Arai, Y. Arai, S. Araki, A. Araya, N. Aritomi, H. Asada, Y. Aso, S. Atsuta *et al.*, KAGRA: 2.5 generation interferometric gravitational wave detector, Nat. Astron. **3**, 35 (2019).

[11] J. E. Kim and G. Carosi, Axions and the strong *CP* problem, Rev. Mod. Phys. **82**, 557 (2010).

[12] K. Choi, S. H. Im, and C. S. Shin, Recent progress in the physics of axions and axion-like particles, Annu. Rev. Nucl. Part. Sci. **71**, 225 (2021).

[13] L. J. Rosenberg and K. A. Van Bibber, Searches for invisible axions, Phys. Rep. **325**, 1 (2000).

[14] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner, and K. A. van Bibber, Experimental searches for the axion and axion-like particles, Annu. Rev. Nucl. Part. Sci. **65**, 485 (2015).

[15] M. Tsang, H. M. Wiseman, and C. M. Caves, Fundamental quantum limit to waveform estimation, Phys. Rev. Lett. **106**, 090401 (2011).

[16] H. Miao, R. X. Adhikari, Y. Ma, B. Pang, and Y. Chen, Towards the fundamental quantum limit of linear measurements of classical signals, Phys. Rev. Lett. **119**, 050801 (2017).

[17] J. W. Gardner, T. Gefen, S. A. Haine, J. J. Hope, and Y. Chen, Achieving the fundamental quantum limit of linear waveform estimation, Phys. Rev. Lett. **132**, 130801 (2024).

[18] 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).

[19] S. Ng, S. Z. Ang, T. A. Wheatley, H. Yonezawa, A. Furusawa, E. H. Huntington, and M. Tsang, Spectrum analysis with quantum dynamical systems, Phys. Rev. A **93**, 042121 (2016).

[20] M. Tsang, Quantum noise spectroscopy as an incoherent imaging problem, Phys. Rev. A **107**, 012611 (2023).

[21] C. W. Gardiner and M. J. Collett, Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation, Phys. Rev. A **31**, 3761 (1985).

[22] R. Kubo, The fluctuation-dissipation theorem, Rep. Prog. Phys. **29**, 255 (1966).

[23] A. Buonanno and Y. Chen, Signal recycled laser-interferometer gravitational-wave detectors as optical springs, Phys. Rev. D **65**, 042001 (2002).

[24] S. L. Danilishin and F. Y. Khalili, Quantum measurement theory in gravitational-wave detectors, Living Rev. Relativ. **15**, 1 (2012).

[25] A. V. Oppenheim and G. C. Verghese, *Signals, Systems & Inference* (Pearson London, London, 2017).

[26] One possible proof of Eq. (3) uses the identity $2\pi\delta(\Omega-\Omega')|_{\Omega=\Omega'} = \lim_{T\to\infty} T$, which follows from $2\pi\delta(a) = \int_{-\infty}^{\infty} \mathrm{d}t\, e^{iat} = \lim_{T\to\infty}\int_0^T \mathrm{d}t\, e^{iat} = \lim_{T\to\infty}\delta_T(a)$, where $\delta_T(a) = (e^{iaT}-1)/(ia)$ is a finite-time approximation to the delta function obeying $\delta_T(0) = T$.

[27] C. Bond, D. Brown, A. Freise, and K. A. Strain, Interferometer techniques for gravitational-wave detection, Living Rev. Relativ. **19**, 1 (2016).

[28] The approximation that $\delta(\Omega-\Omega')|_{\Omega=\Omega'} \approx T/(2\pi)$ for large but finite integration time T provides the correct $\sqrt{2/T}$ normalization in Eq. (10) to satisfy the canonical commutation relations compared to the $\sqrt{1/(\pi T)}$ normalization in Ref. [17], which incorrectly assumes that $\delta(\Omega-\Omega')|_{\Omega=\Omega'} \approx T$.

[29] We consider the variance of the parts $\hat{q}_j(\Omega)$ here since they are nonstationary, such that the power spectral density in Eq. (4) is not defined.

[30] The value of $G_{py}(\Omega)$ in Eq. (14) is equal to twice the value of $G(\Omega)$ from Ref. [70] since we use a different normalization for the input-output relation.

[31] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, Rev. Mod. Phys. **75**, 715 (2003).

[32] W. H. Zurek, in *Quantum Decoherence: Poincaré Seminar 2005* (Birkhäuser Basel, Basel, 2007), pp. 1–31..

[33] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Phys. Rev. Lett. **72**, 3439 (1994).

[34] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).

[35] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, 035005 (2018).

[36] M. Takeoka, M. Ban, and M. Sasaki, Unambiguous quantum-state filtering, Phys. Rev. A **68**, 012307 (2003).

[37] A. Monras, Phase space formalism for quantum estimation of Gaussian states, arxiv:1303.3682 [quant-ph].

[38] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, Phys. Rev. X **6**, 031033 (2016).

[39] S. Zhou and L. Jiang, Modern description of Rayleigh's criterion, Phys. Rev. A **99**, 013808 (2019).

[40] T. Gefen, A. Rotem, and A. Retzker, Overcoming resolution limits with quantum sensing, Nat. Commun. **10**, 4992 (2019).

[41] S. L. Mouradian, N. Glikin, E. Megidish, K.-I. Ellers, and H. Haeffner, Quantum sensing of intermittent stochastic signals, Phys. Rev. A **103**, 032419 (2021).

[42] Y.-x. Liu, Ş. K. Özdemir, A. Miranowicz, and N. Imoto, Kraus representation of a damped harmonic oscillator and its application, Phys. Rev. A **70**, 042308 (2004).
[43] 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).
[44] J. Kolodynski and R. Demkowicz-Dobrzanski, Efficient tools for quantum metrology with uncorrelated noise, New J. Phys. **15**, 073043 (2013).
[45] R. Demkowicz-Dobrzański and L. Maccone, Using entanglement against noise in quantum metrology, Phys. Rev. Lett. **113**, 250801 (2014).
[46] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel, Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector GEO 600, Phys. Rev. A **88**, 041802 (2013).
[47] J. Aasi *et al.*, Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light, Nat. Photonics **7**, 613 (2013).
[48] M. Tse *et al.*, Quantum-enhanced Advanced LIGO detectors in the era of gravitational-wave astronomy, Phys. Rev. Lett. **123**, 231107 (2019).
[49] L. McCuller, C. Whittle, D. Ganapathy, K. Komori, M. Tse, A. Fernandez-Galiana, L. Barsotti, P. Fritschel, M. MacInnis, F. Matichard, K. Mason, N. Mavalvala, R. Mittleman, Haocun Yu, M. E. Zucker, and M. Evanss, Frequency-dependent squeezing for Advanced LIGO, Phys. Rev. Lett. **124**, 171102 (2020).
[50] LIGO O4 Detector Collaboration, D. Ganapathy, W. Jia, M. Nakano, V. Xu, N. Aritomi, T. Cullen, N. Kijbunchoo, S. E. Dwyer, A. Mullavey *et al.*, Broadband quantum enhancement of the LIGO detectors with frequency-dependent squeezing, Phys. Rev. X **13**, 041021 (2023).
[51] W. Górecki, A. Riccardi, and L. Maccone, Quantum metrology of noisy spreading channels, Phys. Rev. Lett. **129**, 240503 (2022).
[52] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).
[53] K. A. Gilmore, M. Affolter, R. J. Lewis-Swan, D. Barberena, E. Jordan, A. M. Rey, and J. J. Bollinger, Quantum-enhanced sensing of displacements and electric fields with two-dimensional trapped-ion crystals, Science **373**, 673 (2021).
[54] X. Deng, S. Li, Z.-J. Chen, Z. Ni, Y. Cai, J. Mai, L. Zhang, P. Zheng, H. Yu, C.-L. Zou *et al.*, Quantum-enhanced metrology with large Fock states, Nat. Phys. **20**, 1874 (2024).
[55] A. Agrawal, A. V. Dixit, T. Roy, S. Chakram, K. He, R. K. Naik, D. I. Schuster, and A. Chou, Stimulated emission of signal photons from dark matter waves, Phys. Rev. Lett. **132**, 140801 (2024).
[56] Q. Zhuang, J. Preskill, and L. Jiang, Distributed quantum sensing enhanced by continuous-variable error correction, New J. Phys. **22**, 022001 (2020).
[57] A. J. Brady, A. Eickbusch, S. Singh, J. Wu, and Q. Zhuang, Advances in bosonic quantum error correction with Gottesman-Kitaev-Preskill codes: Theory, engineering and applications, Prog. Quantum Electron. **93**, 100496 (2024).
[58] S. Krastanov, V. V. Albert, C. Shen, C.-L. Zou, R. W. Heeres, B. Vlastakis, R. J. Schoelkopf, and L. Jiang, Universal control of an oscillator with dispersive coupling to a qubit, Phys. Rev. A **92**, 040303 (2015).
[59] A. Eickbusch, V. Sivak, A. Z. Ding, S. S. Elder, S. R. Jha, J. Venkatraman, B. Royer, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Fast universal control of an oscillator with weak dispersive coupling to a qubit, Nat. Phys. **18**, 1464 (2022).
[60] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New class of quantum error-correcting codes for a bosonic mode, Phys. Rev. X **6**, 031006 (2016).
[61] A. Monras and F. Illuminati, Measurement of damping and temperature: Precision bounds in Gaussian dissipative channels, Phys. Rev. A **83**, 012315 (2011).
[62] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Springer Science & Business Media, Pisa, 2011).
[63] O. Pinel, P. Jian, N. Treps, C. Fabre, and D. Braun, Quantum parameter estimation using general single-mode Gaussian states, Phys. Rev. A **88**, 040102 (2013).
[64] Y. Gao and H. Lee, Bounds on quantum multiple-parameter estimation with Gaussian state, Eur. Phys. J. D **68**, 347 (2014).
[65] H. Nagaoka, in *Asymptotic Theory of Quantum Statistical Inference: Selected Papers* (World Scientific, 2005), pp. 100–112.
[66] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, Efficient computation of the Nagaoka-Hayashi bound for multiparameter estimation with separable measurements, npj Quantum Inf. **7**, 110 (2021).
[67] M. Parniak, S. Borówka, K. Boroszko, W. Wasilewski, K. Banaszek, and R. Demkowicz-Dobrzański, Beating the Rayleigh limit using two-photon interference, Phys. Rev. Lett. **121**, 250503 (2018).
[68] Y. L. Len, Multiparameter estimation for qubit states with collective measurements: A case study, New J. Phys. **24**, 033037 (2022).
[69] If the temporal-basis modes $w(\tau)$ are complex, then their Fourier transforms will have independent positive and negative frequency components, unlike the real basis of cosine and sine functions in Eq. (10). Instead, the real and imaginary parts of each complex temporal mode should be studied, where each part reduces to the canonical noise-estimation problem of a single harmonic oscillator.
[70] S. M. Vermeulen, T. Cullen, D. Grass, I. A. O. MacMillan, A. J. Ramirez, J. Wack, B. Korzh, V. S. H. Lee, K. M. Zurek, C. Stoughton, and L. McCuller, Photon-counting interferometry to detect geontropic space-time fluctuations with GQuEST, Phys. Rev. X **15**, 011034 (2025).
[71] A. V. Gorshkov, A. André, M. D. Lukin, and A. S. Sørensen, Photon storage in Λ-type optically dense atomic media. I. Cavity model, Phys. Rev. A **76**, 033804 (2007).

[72] K. R. Brown, K. M. Dani, D. M. Stamper-Kurn, and K. B. Whaley, Deterministic optical Fock-state generation, Phys. Rev. A **67**, 043818 (2003).

[73] J. M. Geremia, Deterministic and nondestructively verifiable preparation of photon number states, Phys. Rev. Lett. **97**, 073601 (2006).

[74] J. Hastrup and U. L. Andersen, Protocol for generating optical Gottesman-Kitaev-Preskill states with cavity QED, Phys. Rev. Lett. **128**, 170503 (2022).

[75] S. Chen, Y.-A. Chen, T. Strassel, Z.-S. Yuan, B. Zhao, J. Schmiedmayer, and J.-W. Pan, Deterministic and storable single-photon source based on a quantum memory, Phys. Rev. Lett. **97**, 173004 (2006).

[76] S. E. Thomas, L. Wagner, R. Joos, R. Sittig, C. Nawrath, P. Burdekin, I. Maillette de Buy Wenniger, M. J. Rasiah, T. Huber-Loyola, S. Sagona-Stophel, S. Höfling, M. Jetter, P. Michler, I. A. Walmsley, S. L. Portalupi, and P. M. Ledingham, Deterministic storage and retrieval of telecom light from a quantum dot single-photon source interfaced with an atomic quantum memory, Sci. Adv. **10**, eadi7346 (2024).

[77] P. Farrera, G. Heinze, B. Albrecht, M. Ho, M. Chávez, C. Teo, N. Sangouard, and H. de Riedmatten, Generation of single photons with highly tunable wave shape from a cold atomic ensemble, Nat. Commun. **7**, 13556 (2016).

[78] O. Katz and O. Firstenberg, Light storage for one second in room-temperature alkali vapor, Nat. Commun. **9**, 2074 (2018).

[79] K. Bergmann, H. Theuer, and B. W. Shore, Coherent population transfer among quantum states of atoms and molecules, Rev. Mod. Phys. **70**, 1003 (1998).

[80] N. V. Vitanov, A. A. Rangelov, B. W. Shore, and K. Bergmann, Stimulated Raman adiabatic passage in physics, chemistry, and beyond, Rev. Mod. Phys. **89**, 015006 (2017).

[81] I. Bloch, Ultracold quantum gases in optical lattices, Nat. Phys. **1**, 23 (2005).

[82] A. M. Kaufman and K.-K. Ni, Quantum science with optical tweezer arrays of ultracold atoms and molecules, Nat. Phys. **17**, 1324 (2021).

[83] A. Reiserer and G. Rempe, Cavity-based quantum networks with single atoms and optical photons, Rev. Mod. Phys. **87**, 1379 (2015).

[84] K. Hammerer, A. S. Sørensen, and E. S. Polzik, Quantum interface between light and atomic ensembles, Rev. Mod. Phys. **82**, 1041 (2010).

[85] K. Heshami, D. G. England, P. C. Humphreys, P. J. Bustard, V. M. Acosta, J. Nunn, and B. J. Sussman, Quantum memories: Emerging applications and recent advances, J. Mod. Opt. **63**, 2005 (2016).

[86] S.-H. Wei, B. Jing, X.-Y. Zhang, J.-Y. Liao, C.-Z. Yuan, B.-Y. Fan, C. Lyu, D.-L. Zhou, Y. Wang, G.-W. Deng *et al.*, Towards real-world quantum networks: A review, Laser Photonics Rev. **16**, 2100219 (2022).

[87] R. Acharya, S. Brebels, A. Grill, J. Verjauw, Ts. Ivanov, D. Perez Lozano, D. Wan, J. Van Damme, A. M. Vadiraj, M. Mongillo, B. Govoreanu, J. Craninckx, I. P. Radu, K. De Greve, G. Gielen, F. Catthoor, and A. Potočnik, Multiplexed superconducting qubit control at millikelvin temperatures with a low-power cryo-CMOS multiplexer, Nat. Electron. **6**, 900 (2023).

[88] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, Phys. Rev. A **69**, 062320 (2004).

[89] L. Chen, H.-X. Li, Y. Lu, C. W. Warren, C. J. Križan, S. Kosen, M. Rommel, S. Ahmed, A. Osman, J. Biznárová, A. F. Roudsari, B. Lienhard, M. Caputo, K. Grigoras, L. Grönberg, J. Govenius, A. F. Kockum, P. Delsing, J. Bylander, and G. Tancredi, Transmon qubit readout fidelity at the threshold for quantum error correction without a quantum-limited amplifier, npj Quantum Inf. **9**, 26 (2023).

[90] M. H. Devoret and R. J. Schoelkopf, Superconducting circuits for quantum information: An outlook, Science **339**, 1169 (2013).

[91] D. P. DiVincenzo, The physical implementation of quantum computation, Fortschritte Phys. **48**, 771 (2000).

[92] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret, Quantum error correction of a qubit encoded in grid states of an oscillator, Nature **584**, 368 (2020).

[93] A. A. Diringer, E. Blumenthal, A. Grinberg, L. Jiang, and S. Hacohen-Gourgy, Conditional-NOT displacement: Fast multioscillator control with a single qubit, Phys. Rev. X **14**, 011055 (2024).

[94] R. Moretti, H. A. Corti, D. Labranca, F. Ahrens, G. Avallone, and D. Babusci, Design and simulation of a transmon qubit chip for axion detection, IEEE Trans. Appl. Supercond. **34**, 1 (2024).

[95] S. K. Lamoreaux, K. A. van Bibber, K. W. Lehnert, and G. Carosi, Analysis of single-photon and linear amplifier detectors for microwave cavity dark matter axion searches, Phys. Rev. D **88**, 035020 (2013).

[96] If measuring signals below the classical noise floor, then we expect the regime $\eta \gg 4\sigma_C^2$ to be also relevant; see the discussion surrounding Eq. (39) for the one-dimensional case.

[97] A. I. Renzini, B. Goncharov, A. C. Jenkins, and P. M. Meyers, Stochastic gravitational-wave backgrounds: Current detection efforts and future prospects, Galaxies **10**, 34 (2022).

[98] A. Buikema *et al.*, Sensitivity and performance of the Advanced LIGO detectors in the third observing run, Phys. Rev. D **102**, 062003 (2020).

[99] This argument assumes a uniform source density, which is valid for the current generation of gravitational-wave observatories. Future observatories are proposed to have much further reach [113,114], such that this assumption is no longer valid and a more detailed astrophysical model is required.

[100] K. Chatziioannou, J. A. Clark, A. Bauswein, M. Millhouse, T. B. Littenberg, and N. Cornish, Inferring the post-merger gravitational wave emission from binary neutron star coalescences, Phys. Rev. D **96**, 124035 (2017).

[101] A. W. Criswell, J. Miller, N. Woldemariam, T. Soultanis, A. Bauswein, K. Chatziioannou, M. W. Coughlin, G. Jones, and V. Mandic, Hierarchical Bayesian method for constraining the neutron star equation of state with

an ensemble of binary neutron star postmerger remnants, Phys. Rev. D **107**, 043021 (2023).

[102] A. Sasli, N. Karnesis, and N. Stergioulas, Exploring the potential for detecting rotational instabilities in binary neutron star merger remnants with gravitational wave detectors, Phys. Rev. D **109**, 043045 (2024).

[103] In the deterministic event-stacking case [100–102], there are threshold effects where the CCRB cannot be saturated [153,154]. This behavior arises from the nonlinear process of frequency estimation at a low signal-to-noise ratio and can be expressed as the discrepancy between the classical Cramér-Rao and Barankin bounds [155,156]. Since finding the frequency of excess power is a stochastic estimation problem, studying the quantum analogs of the Barankin bound [157] may be necessary for understanding the behavior here at low signal-to-noise ratios.

[104] Finding the optimal measurement for fixed finite M would require instead calculating the Nagaoka-Hayashi bound [65,66].

[105] R. Cameron, G. Cantatore, A. Melissinos, G. Ruoso, Y. Semertzidis, H. Halama, D. Lazarus, A. Prodell, F. Nezrick, C. Rizzo *et al.*, Search for nearly massless, weakly coupled particles by optical techniques, Phys. Rev. D **47**, 3707 (1993).

[106] N. Du, N. Force, R. Khatiwada, E. Lentz, R. Ottens, L. Rosenberg, G. Rybka, G. Carosi, N. Woollett, and D. Bowring *et al.*, Search for invisible axion dark matter with the axion dark matter experiment, Phys. Rev. Lett. **120**, 151301 (2018).

[107] A. V. Dixit, S. Chakram, K. He, A. Agrawal, R. K. Naik, D. I. Schuster, and A. Chou, Searching for dark matter with a superconducting qubit, Phys. Rev. Lett. **126**, 141302 (2021).

[108] H. Shi and Q. Zhuang, Ultimate precision limit of noise sensing and dark matter search, npj Quantum Inf. **9**, 27 (2023).

[109] H. Zheng, M. Silveri, R. Brierley, S. Girvin, and K. Lehnert, Accelerating dark-matter axion searches with quantum measurement technology, arxiv:1607.02529 [hep-ph].

[110] K. M. Backes *et al.*, A quantum enhanced search for dark matter axions, Nature **590**, 238 (2021).

[111] C. M. Caves, Quantum-mechanical noise in an interferometer, Phys. Rev. D **23**, 1693 (1981).

[112] L. McCuller, C. Whittle, D. Ganapathy, K. Komori, M. Tse, A. Fernandez-Galiana, L. Barsotti, P. Fritschel, M. MacInnis, and F. Matichard, Frequency-dependent squeezing for Advanced LIGO, Phys. Rev. Lett. **124**, 171102 (2020).

[113] D. Reitze *et al.*, Astro2020: Decadal Survey on Astronomy and Astrophysics, APC white papers, in *Bulletin of the American Astronomical Society* (2019), Vol. 51.

[114] M. Maggiore, C. van den Broeck, N. Bartolo, E. Belgacem, D. Bertacca, M. A. Bizouard, M. Branchesi, S. Clesse, S. Foffa, J. García-Bellido, S. Grimm, J. Harms, T. Hinderer, S. Matarrese, C. Palomba, M. Peloso, A. Ricciardone, and M. Sakellariadou, Science case for the Einstein telescope, J. Cosmol. Astropart. Phys. **2020**, 050 (2020).

[115] R. D. Gill and B. Y. Levit, Applications of the van Trees inequality: A Bayesian Cramér-Rao bound, Bernoulli **1**, 59 (1995).

[116] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part IV* (John Wiley & Sons, Incorporated, New York, 2002).

[117] A. Bednorz and W. Belzig, Fourth moments reveal the negativity of the Wigner function, Phys. Rev. A **83**, 052113 (2011).

[118] R. Howl, V. Vedral, D. Naik, M. Christodoulou, C. Rovelli, and A. Iyer, Non-Gaussianity as a signature of a quantum theory of gravity, PRX Quantum **2**, 010325 (2021).

[119] S. A. Haine, Searching for signatures of quantum gravity in quantum gases, New J. Phys. **23**, 033020 (2021).

[120] Z. Mehdi, J. J. Hope, and S. A. Haine, Signatures of quantum gravity in the gravitational self-interaction of photons, Phys. Rev. Lett. **130**, 240203 (2023).

[121] A. Franzen, Componentlibrary, 2009. http://www.gwoptics.org/ComponentLibrary/.

[122] Wolfram Research, Inc., *Mathematica* 8.0 (2010). https://www.wolfram.com.

[123] G. Van Rossum and F. L. Drake, PYTHON *3 Reference Manual* (CreateSpace, Scotts Valley, California, 2009).

[124] F. Pérez and B. E. Granger, IPython: A system for interactive scientific computing, Comput. Sci. Eng. **9**, 21 (2007).

[125] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing, and Jupyter Development Team, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas* (IOS Press, 2016), pp. 87–90.

[126] C. R. Harris *et al.*, Array programming with NumPy, Nature **585**, 357 (2020).

[127] J. D. Hunter, MATPLOTLIB: A 2D graphics environment, Comput. Sci. Eng. **9**, 90 (2007).

[128] Wes McKinney, in *Proceedings of the 9th* PYTHON *in Science Conference* (SciPy, Austin, Texas, 2010), pp. 56–61.

[129] P. Virtanen *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in PYTHON, Nat. Methods **17**, 261 (2020).

[130] J. Johansson, P. Nation, and F. Nori, QuTiP 2: A PYTHON framework for the dynamics of open quantum systems, Comput. Phys. Commun. **184**, 1234 (2013).

[131] N. Killoran, J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook, STRAWBERRY FIELDS: A software platform for photonic quantum computing, Quantum **3**, 129 (2019).

[132] J. W. Gardner, *pleasantPheasant* (2023). https://git.ligo.org/jameswalter.gardner/pleasantpheasant.

[133] J. Williamson, On the algebraic problem concerning the normal forms of linear dynamical systems, Am. J. Math. **58**, 141 (1936).

[134] S. A. Haine, Using interaction-based readouts to approach the ultimate limit of detection-noise robustness for quantum-enhanced metrology in collective spin systems, Phys. Rev. A **98**, 030303 (2018).

[135] Y. L. Len, T. Gefen, A. Retzker, and J. Kołodyński, Quantum metrology with imperfect measurements, Nat. Commun. **13**, 6971 (2022).

[136] S. Zhou, S. Michalakis, and T. Gefen, Optimal protocols for quantum metrology with noisy measurements, PRX Quantum **4**, 040305 (2023).
[137] A. Uhlmann, The "transition probability" in the state space of a $*$-algebra, Rep. Math. Phys. **9**, 273 (1976).
[138] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Generalized limits for single-parameter quantum estimation, Phys. Rev. Lett. **98**, 090401 (2007).
[139] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, J. Phys. A **41**, 255304 (2008).
[140] S. Kawamura and Y. Chen, Displacement-noise-free gravitational-wave detection, Phys. Rev. Lett. **93**, 211103 (2004).
[141] Y. Chen, A. Pai, K. Somiya, S. Kawamura, S. Sato, K. Kokeyama, R. L. Ward, K. Goda, and E. E. Mikhailov, Interferometers for displacement-noise-free gravitational-wave detection, Phys. Rev. Lett. **97**, 151103 (2006).
[142] T. Gefen, R. Tarafder, R. X. Adhikari, and Y. Chen, Quantum precision limits of displacement noise-free interferometers, Phys. Rev. Lett. **132**, 020801 (2024).
[143] I. Tzitrin, J. E. Bourassa, N. C. Menicucci, and K. K. Sabapathy, Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes, Phys. Rev. A **101**, 032315 (2020).
[144] J. E. Bourassa, N. Quesada, I. Tzitrin, A. Száva, T. Isacsson, J. Izaac, K. K. Sabapathy, G. Dauphinais, and I. Dhand, Fast simulation of bosonic qubits via Gaussian functions in phase space, PRX Quantum **2**, 040315 (2021).
[145] K. Macieszczak, Quantum Fisher information: Variational principle and simple iterative algorithm for its efficient computation, arxiv:1312.1356 [quant-ph].
[146] Á. Lukács, R. Trényi, T. Vértesi, and G. Tóth, Optimizing local Hamiltonians for the best metrological performance, arXiv:2206.02820.
[147] J. Gorski, F. Pfeuffer, and K. Klamroth, Biconvex sets and optimization with biconvex functions: A survey and extensions, Math. Methods Oper. Res. **66**, 373 (2007).
[148] C. A. Floudas and V. Visweswaran, A global optimization algorithm (GOP) for certain classes of nonconvex NLPs—I. Theory, Comput. Chem. Eng. **14**, 1397 (1990).
[149] C. A. Floudas, *Deterministic Global Optimization: Theory, Methods and Applications* (Springer Science & Business Media, New York, 2013), Vol. 37.
[150] K. Noh, V. V. Albert, and L. Jiang, Quantum capacity bounds of Gaussian thermal loss channels and achievable rates with Gottesman-Kitaev-Preskill codes, IEEE Trans. Inf. Theory **65**, 2563 (2018).
[151] G. Tóth and T. Vértesi, Quantum states with a positive partial transpose are useful for metrology, Phys. Rev. Lett. **120**, 020506 (2018).
[152] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, Bayesian quantum frequency estimation in presence of collective dephasing, New J. Phys. **16**, 113002 (2014).
[153] D. Rife and R. Boorstyn, Single tone parameter estimation from discrete-time observations, IEEE Trans. Inf. Theory **20**, 591 (1974).
[154] A. Steinhardt and C. Bretherton, in *ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing* (Tampa, FL, USA, 1985), Vol. 10, pp. 1273–1276.
[155] L. Knockaert, The Barankin bound and threshold behavior in frequency estimation, IEEE Trans. Signal Process. **45**, 2398 (1997).
[156] E. Chaumette, Jé. Galy, A. Quinlan, and P. Larzabal, A new Barankin bound approximation for the prediction of the threshold region performance of maximum likelihood estimators, IEEE Trans. Signal Process. **56**, 5319 (2008).
[157] M. Gessner and A. Smerzi, Hierarchies of frequentist bounds for quantum metrology: From Cramér-Rao to Barankin, Phys. Rev. Lett. **130**, 260801 (2023).