

# Stochastic Heisenberg limit: Optimal estimation of a fluctuating phase

Dominic W. Berry<sup>1</sup>, Michael J. W. Hall<sup>2</sup>, and Howard M. Wiseman<sup>2</sup>

<sup>1</sup>*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*

<sup>2</sup>*Centre for Quantum Computation and Communication Technology (Australian Research Council),  
Centre for Quantum Dynamics, Griffith University, Brisbane, QLD 4111, Australia*

The ultimate limits to estimating a fluctuating phase imposed on an optical beam can be found using the recently derived continuous quantum Cramér-Rao bound. For Gaussian stationary statistics, and a phase spectrum scaling asymptotically as  $\omega^{-p}$  with  $p > 1$ , the minimum mean-square error in any (single-time) phase estimate scales as  $\mathcal{N}^{-2(p-1)/(p+1)}$ , where  $\mathcal{N}$  is the photon flux. This gives the usual Heisenberg limit for a constant phase (as the limit  $p \rightarrow \infty$ ) and provides a stochastic Heisenberg limit for fluctuating phases. For  $p = 2$  (Brownian motion), this limit can be attained by phase tracking.

PACS numbers: 42.50.St, 03.65.Ta, 06.20.Dk

Estimating the phase imposed on an optical beam, by nature or by an agent, is a key task in metrology and communication respectively. One case of broad relevance is that where the phase varies stochastically in time over a wide range [1–8]. It is only very recently that it has been possible to experimentally demonstrate the quantum enhancement (by a constant factor) of the estimation of such a strongly fluctuating phase, using nonclassical (squeezed) light and homodyne detection with adaptive phase tracking [1, 7].

Adaptive phase tracking is a sophisticated measurement technique whereby the phase of the local oscillator (necessary for homodyne detection) is continuously changed in time to follow an estimate of the true phase [1–7]. This enables the phase quadrature of the beam to be monitored at all times, to a good approximation, maximizing the phase information obtained. Previously it has been calculated that phase tracking with squeezed light would enable an imposed phase to be estimated with a mean square error (MSE) scaling as  $\mathcal{N}^{-2/3}$  [2–4]. In contrast, for coherent states (no squeezing) only a  $\mathcal{N}^{-1/2}$  scaling can be achieved [2–4]. Here  $\mathcal{N}$  is the mean flux (photons per second) in the beam, and the imposed phase is modeled by Brownian motion.

While experiments in optical phase tracking have not yet demonstrated an improvement over the coherent state scaling of  $\mathcal{N}^{-1/2}$ , the possibility of doing so in the near future raises pressing theoretical questions: is the MSE scaling of  $\mathcal{N}^{-2/3}$ , derived assuming adaptive estimation [2–4], the best possible? If not, what is the ultimate limit to estimating a fluctuating phase and how can it be achieved?

For measurement of a *constant* phase, the fundamental bound is the Heisenberg limit [9, 10]: a phase estimate MSE scaling as  $\langle N \rangle^{-2}$ , where  $\langle N \rangle$  is the mean number of photons per estimate. This a quadratic improvement over the  $\langle N \rangle^{-1}$  scaling achievable using coherent states (the standard quantum limit, or SQL) [9, 10]. Hence, if quantum mechanics similarly allowed a quadratic improvement in the case of a *fluctuating* phase, the corre-

sponding fundamental limit for the MSE would scale as  $\mathcal{N}^{-1}$ .

Contrary to this intuition, we prove in this paper, with only weak assumptions, that the fundamental bound to estimating Brownian phase fluctuations is a MSE scaling as  $\mathcal{N}^{-2/3}$ . This establishes that adaptive phase tracking can be a very effective measurement technique for this problem, giving an uncertainty at most a constant factor greater than the minimum allowed by quantum mechanics (under our assumptions). This  $\mathcal{N}^{-2/3}$  scaling for Brownian fluctuations is just a special case of our general stochastic Heisenberg limit, which allows for any inverse power-law describing the phase fluctuation spectrum at high frequencies, and which also yields the constant-phase Heisenberg limit as a special case.

This paper is organised as follows. First we derive the general stochastic Heisenberg limit, and the stochastic SQL. Next we specialize to the scenario of Ref. [1]: a squeezed beam comprising the output of an optical parametric oscillator (OPO) with an added mean field, and a phase varying like damped Brownian motion. We consider the ultimate limit, and find the same scaling as in the general case, but with an explicit constant of proportionality, consistent with the numerics of Ref. [4].

## GENERAL PROOF

Our result applies to the situation of a continuous beam [a one-dimensional quantum field  $b(t)$ ], on which there is an imposed phase  $\varphi(t)$ . We require only three conditions:

1. The statistics of the field quadratures and imposed phase are *stationary*.
2. The statistics of the field quadratures and imposed phase are *Gaussian* and *time-symmetric*.
3. The phase spectrum scales as  $|\omega|^{-p}$  for large  $|\omega|$ , for some  $p > 1$ .

We now explain these conditions in more detail. The instantaneous creation operator  $b^\dagger(t)$  of the beam obeys  $[b(t), b^\dagger(t')] = \delta(t - t')$ , and  $\langle b^\dagger(t)b(t) \rangle = \mathcal{N}$  is the photon flux [10]. The stationarity of its statistics, and those of  $\varphi(t)$ , means that one-time expectation values are constant, and two-time correlations depend only on the time difference.

The assumed Gaussian statistics mean there is a Gaussian Wigner functional for the quadratures [10]

$$X(t) := b^\dagger(t) + b(t), \quad Y(t) := i[b^\dagger(t) - b(t)], \quad (1)$$

and a Gaussian distribution for the phase  $\varphi(t)$ . For a single Gaussian variable, such as  $\varphi(t)$ , the autocorrelation function is automatically time-symmetric. However, the beam quadratures  $X(t)$  and  $Y(t')$  may be correlated, and our derivation below requires that this cross-correlation be time-symmetric (i.e. invariant under  $t \leftrightarrow t'$ ).

The third condition allows for a vast range of phase fluctuation models. For  $p = 2$  it means that at short times the fluctuations are like Wiener noise [10] (Brownian motion), which has the spectrum  $\kappa/\omega^2$ , where  $\kappa$  is a constant with units of frequency. More generally, we take the scaling constant to be  $\kappa^{p-1}$ , so  $\kappa$  still has units of frequency. In the limit  $p \rightarrow \infty$  the phase is effectively constant.

In Ref. [11], a continuous form of the quantum Cramér-Rao inequality was derived, giving a lower bound on the MSE of any unbiased estimate,  $\hat{\varphi}(t)$ , of a time-varying parameter  $\varphi(t)$ ,

$$\langle [\hat{\varphi}(t) - \varphi(t)]^2 \rangle \geq F^{-1}(t, t). \quad (2)$$

Here  $F(t, t')$  is the Fisher information matrix (with continuous indices  $t$  and  $t'$ ) of the phase of the beam, given by a sum of quantum and classical contributions

$$F(t, t') := F^{(Q)}(t, t') + F^{(C)}(t, t'), \quad (3)$$

and the (matrix) inverse in Eq. (2) is defined by

$$\int ds F^{-1}(t, s) F(s, t') = \delta(t - t'). \quad (4)$$

For the case where  $\varphi(t)$  is an imposed phase,

$$F^{(Q)}(t, t') := 4 \langle \Delta n(t) \Delta n(t') \rangle, \quad (5)$$

$$F^{(C)}(t, t') := \int D\varphi P[\varphi] \frac{\delta \ln P[\varphi]}{\delta \varphi(t)} \frac{\delta \ln P[\varphi]}{\delta \varphi(t')}. \quad (6)$$

In the above,  $\int D\varphi \dots$  denotes an integral over all possible functions  $\varphi(t)$ , the functional  $P[\varphi]$  gives the prior weight for each function, and  $\delta/\delta\varphi(t)$  is a functional derivative. Also,  $\Delta n(t) = n(t) - \langle n(t) \rangle$ , where  $n(t) := b^\dagger(t)b(t)$  is the generator of the phase shifts, the photon flux operator. Because of the stationarity condition, all quantities dependent on two times  $t$  and  $t'$  are functions only of  $t - t'$ . We will express these quantities explicitly

as functions of  $t - t'$  from here on. In particular, the lower bound in Eq. (2) will be denoted by  $F^{-1}(0)$ .

To determine  $F^{-1}(0)$ , substitute Eq. (3) into Eq. (4) and take the Fourier transform, to give

$$\tilde{F}^{-1}(\omega) = \frac{1}{\tilde{F}^{(C)}(\omega) + \tilde{F}^{(Q)}(\omega)}, \quad (7)$$

for the Fourier transform of  $F^{-1}(t - t')$ . The value of  $F^{-1}(0)$  is then obtained by integrating this over  $\omega$ . Our aim is to determine the minimum possible scaling of this value with the photon flux  $\mathcal{N} = \langle n(t) \rangle$ .

As we assume the phase fluctuations are Gaussian, we have  $F^{(C)}(t - t') = \Sigma^{-1}(t - t')$ , with  $\Sigma(t - t') := \langle \varphi(t)\varphi(t') \rangle - \langle \varphi \rangle^2$  [12]. For the case  $\tilde{\Sigma}(\omega) = \kappa^{p-1}/|\omega|^p$ ,

$$\tilde{F}^{-1}(\omega) = \frac{\kappa^{p-1}}{|\omega|^p + \kappa^{p-1} \tilde{F}^{(Q)}(\omega)}. \quad (8)$$

More generally, we can obtain the result below with the weaker requirement that the phase spectrum approaches this scaling at high frequencies, i.e.,  $\tilde{\Sigma}(\omega) = \Omega(\kappa^{p-1}/|\omega|^p)$  as per condition 3 (see Appendix D).

Next we consider the quantity  $F^{(Q)}(t - t')$ ; this may be simplified to (see Appendix A)

$$F^{(Q)}(t - t') = 4\mathcal{N}\delta(t - t') + f(t - t') - g(t - t'), \quad (9)$$

where, in terms of the quadrature operators (1),

$$\begin{aligned} f(t - t') &:= \frac{1}{2} [ \langle : X(t)X(t') : \rangle + \langle : Y(t)Y(t') : \rangle ]^2 \\ g(t - t') &:= \langle : X(t)X(t') : \rangle \langle : Y(t)Y(t') : \rangle \\ &\quad - \frac{1}{2} ( \langle : X(t)Y(t') : \rangle^2 + \langle : Y(t)X(t') : \rangle^2 ) \\ &\quad + \frac{1}{2} ( \langle X \rangle^2 + \langle Y \rangle^2 )^2. \end{aligned} \quad (11)$$

Thus, from Eq. (8), we obtain

$$\tilde{F}^{-1}(\omega) = \frac{\kappa^{p-1}}{|\omega|^p + \kappa^{p-1} [4\mathcal{N} + \tilde{f}(\omega) - \tilde{g}(\omega)]}. \quad (12)$$

The photon flux can be written as

$$\mathcal{N} = \frac{1}{4} ( \langle : X(t)X(t) : \rangle + \langle : Y(t)Y(t) : \rangle ), \quad (13)$$

and therefore  $f(0) = 8\mathcal{N}^2$ . In addition, using a spectral uncertainty principle and the assumption of time-symmetric correlations, it can be shown that  $\mathcal{N}/4 \geq -\tilde{g}(\omega)$  and  $\tilde{f}(\omega) \geq 0$  (see Appendix B). Since it is easily shown that  $F^{(Q)}(t - t')$  is a positive-definite function, by Bochner's theorem  $\tilde{F}^{(Q)}(\omega) \geq 0$  [13], and thus the denominator in Eq. (12) is positive. Consequently, replacing  $-\tilde{g}(\omega)$  with  $\mathcal{N}/4$  can only decrease the right-hand side; that is, with  $\zeta = 17/4$ ,

$$\tilde{F}^{-1}(\omega) \geq \frac{\kappa^{p-1}}{|\omega|^p + \kappa^{p-1} [ \zeta \mathcal{N} + \tilde{f}(\omega) ]}. \quad (14)$$

Next, from the fact that  $f(0) = 8\mathcal{N}^2$ , the integral of  $\tilde{f}(\omega)$  is  $\mathcal{I} = 16\pi\mathcal{N}^2$ . This means that  $\tilde{f}(\omega)$  cannot be larger than  $\mu$  over a range greater than  $\mathcal{I}/\mu$ . To place a lower bound on  $F^{-1}(0)$ , when integrating Eq. (14) we may first omit the range of integration where  $\tilde{f}(\omega) > \mu$ , and replace  $\tilde{f}(\omega)$  by  $\mu$  over the remaining portion. Second, we can assume that the range of integration omitted is for the smallest values of  $\omega$  because that can only further reduce the value of the integral. Since this range can be at most  $\mathcal{I}/\mu$  in length, this yields

$$\begin{aligned} F^{-1}(0) &\geq \frac{1}{\pi} \int_{\mathcal{I}/2\mu}^{\infty} \frac{\kappa^{p-1} d\omega}{\omega^p + \kappa^{p-1} (\zeta\mathcal{N} + \mu)} \\ &\geq \frac{1}{2\pi} \int_0^{\infty} \frac{\kappa^{p-1} d\omega}{\omega^p + \kappa^{p-1} (\zeta\mathcal{N} + \mu)} \\ &\geq \frac{\kappa^{p-1}}{2\pi[\kappa^{p-1}(\zeta\mathcal{N} + \mu)]^{1-1/p}}, \end{aligned} \quad (15)$$

where the inequality on the second line holds for  $(\mathcal{I}/\mu)^p \leq \kappa^{p-1} (\zeta\mathcal{N} + \mu)$  (see Appendix C).

To obtain the strongest lower bound on  $F^{-1}(0)$  we consider the smallest value of  $\mu$  that we can take such that  $(\mathcal{I}/\mu)^p \leq \kappa^{p-1} (\zeta\mathcal{N} + \mu)$ . This value is  $\mu = \Theta(\mathcal{N}(\mathcal{N}/\kappa)^{(p-1)/(p+1)})$ . We consider scaling for large  $\mathcal{N}/\kappa$ , in which case  $\mu \gg \mathcal{N}$ , and  $\zeta\mathcal{N}$  can be ignored. Equations (2) and (15) thus yield our main result, the lower bound scaling for the MSE,

$$\langle [\hat{\varphi}(t) - \varphi(t)]^2 \rangle = \Omega \left( (\kappa/\mathcal{N})^{2(p-1)/(p+1)} \right). \quad (16)$$

Note this scaling cannot be achieved by a coherent-state beam, for which  $F^{(Q)}(t-t') = 4\mathcal{N}\delta(t-t')$ . It is easy to show, for this case, by taking  $f = g = 0$  in Eq. (12), that

$$\langle [\hat{\varphi}(t) - \varphi(t)]^2 \rangle_{\text{SQL}} = \Omega \left( (\kappa/\mathcal{N})^{(p-1)/p} \right). \quad (17)$$

which we call the stochastic SQL scaling.

In the case of Wiener phase fluctuations ( $p = 2$ ), the stochastic Heisenberg scaling is  $(\kappa/\mathcal{N})^{2/3}$ . A simplified analysis in Ref. [2] found that adaptive homodyne measurements can yield this scaling, but it took into account neither the photon flux due to the squeezing nor the information in the photocurrent noise. A more complete analysis, taking both of these terms into account, was performed in Ref. [4] (correcting an error in the analysis of Ref. [3]). This verified this scaling of  $(\kappa/\mathcal{N})^{2/3}$  for the MSE. That is, the lower bound in Eq. (16) is attainable by adaptive measurements for  $p = 2$ .

The limit  $p \rightarrow \infty$  gives a very slowly varying phase. In that case Eq. (16) gives the expected constant-phase Heisenberg limit scaling  $\langle [\hat{\varphi}(t) - \varphi(t)]^2 \rangle = \Omega(\langle \mathcal{N} \rangle^{-2})$ . Similarly, Eq. (17) gives the expected SQL scaling  $\langle [\hat{\varphi}(t) - \varphi(t)]^2 \rangle_{\text{SQL}} = \Omega(\langle \mathcal{N} \rangle^{-1})$ . For all other  $p$ , the quantum enhancement is less than quadratic, and as  $p \rightarrow 1$  there is no quantum advantage.

## OPO SQUEEZING AND OU FLUCTUATIONS

Next we specialise to the model of squeezing used in Refs. [1–3]—a coherent field of real amplitude  $\alpha$  added to an Optical Parametric Oscillator (OPO) output—and to phase fluctuations modeled by Ornstein-Uhlenbeck (OU) noise, with  $\tilde{\Sigma}(\omega) = \kappa/(\lambda^2 + \omega^2)$  as in Ref. [1]. Asymptotically this is identical to the Wiener phase spectrum ( $p = 2$ ) analysed above. For this beam we have  $\langle X \rangle = 2\alpha$ ,  $\langle Y \rangle = 0$ , so Eq. (9) becomes

$$\begin{aligned} F^{(Q)}(t-t') &= 4\mathcal{N}\delta(t-t') + 4\alpha^2 T_+(t-t') \\ &\quad + \{ [T_+(t-t')]^2 + [T_-(t-t')]^2 \} / 2, \end{aligned} \quad (18)$$

where  $T_{\pm}(t-t')$  are the normally ordered correlation functions for the quadrature fluctuations [1, 15]

$$T_{\pm}(t-t') = \langle : \Delta Q_{\pm}(t) \Delta Q_{\pm}(t') : \rangle, \quad (19)$$

where  $Q_+ := X$  and  $Q_- := Y$ . They are given by

$$T_{\pm}(t-t') = (R_{\pm} - 1) \frac{(1 \mp x)\gamma}{4} e^{-(1 \mp x)\gamma|t-t'|/2}. \quad (20)$$

In Eq. (20),  $R_{\pm}$  are the antisqueezing and squeezing levels, respectively, at the center frequency. For an OPO,  $\gamma$  is the cavity's decay rate [16] and  $x \in [0, 1)$  is the normalized pump amplitude. In terms of these quantities, the total photon flux is

$$\mathcal{N} = \alpha^2 + \frac{\gamma}{16} [(R_+ - 1)(1 - x) + (R_- - 1)(1 + x)]. \quad (21)$$

Substituting these expressions in Eq. (18) and taking the Fourier transform yields

$$\begin{aligned} \tilde{F}^{(Q)}(\omega) &= 4\mathcal{N} + 4\alpha^2 (R_+ - 1) \frac{(1-x)^2 \gamma^2}{(1-x)^2 \gamma^2 + 4\omega^2} \\ &\quad + \frac{\gamma^3}{16} \left[ \frac{(R_+ - 1)^2 (1-x)^3}{(1-x)^2 \gamma^2 + \omega^2} + \frac{(R_- - 1)^2 (1+x)^3}{(1+x)^2 \gamma^2 + \omega^2} \right]. \end{aligned} \quad (22)$$

For a coherent state ( $R_+ = R_- = 1$ ), we would just have  $\tilde{F}^{(Q)}(\omega) = 4\mathcal{N}$ , and we would obtain  $\kappa/(2\sqrt{4\mathcal{N}\kappa + \lambda^2})$  as the lower bound on the MSE. This coherent state limit was first derived by Tsang *et al.* (see Eq. (4.5) in Ref. [5]) and scales asymptotically as  $(\kappa/\mathcal{N})^{1/2}$  as expected [2].

## Comparison with the Science experiment

It seems impossible to obtain an exact analytical solution for  $F^{-1}(0)$  in the case of general OPO squeezing. However, a useful approximation is to just include the terms in Eq. (22) that represent information available from the mean field; that is, those terms proportional to  $\alpha^2$ . As in the theory of Ref. [2], the estimation performed in Ref. [1] only used the signal from the mean field, so this approximation is relevant to those works. As in those

works we also express our results in terms of  $\alpha^2$ , rather than  $\mathcal{N}$ . Then we find

$$F^{-1}(0) = \frac{(\kappa/2) (1 + g^2/\sqrt{\Xi_+\Xi_-})}{\sqrt{\Xi_+} + \sqrt{\Xi_-}}, \quad (23)$$

where

$$\Xi_{\pm} = \frac{1}{2} \left( 4\alpha^2\kappa + \lambda^2 + g^2 \pm \sqrt{(4\alpha^2\kappa + \lambda^2 - g^2)^2 - 4d} \right), \quad (24)$$

with  $g = (1-x)\gamma/2$  and  $d = 4\kappa\alpha^2(R_+ - 1)g^2$ .

With further simplification this bound on the MSE is comparable to the results given for adaptive measurements on squeezed states in Ref. [1]. First we note that a mixed squeezed state described by  $R_{\pm}$ ,  $\gamma$  and  $x$  can be assumed to be a combination of a pure squeezed state and classical amplitude noise (see Appendix E), where the pure squeezed state is described by  $R_-^Q = R_-$ ,  $R_+^Q = 1/R_-$ ,  $x^Q = (\sqrt{R_+} - 1)/(\sqrt{R_+} + 1)$ , and  $\gamma^Q = \gamma(1+x)/(1+x^Q)$ . Then one can determine  $F^{(Q)}$  from the parameters for the pure squeezed state.

Using this approach, in the limit of large bandwidth,  $\gamma \rightarrow \infty$ , we obtain  $\kappa/(2\sqrt{4\alpha^2 R_- \kappa + \lambda^2})$  as the lower bound on the MSE. This expression is that shown as trace (iii) in Fig. 3 of Ref. [1], derived from Eq. (3) of Ref. [1] by taking  $\sigma_f \rightarrow 0$  (the limit of perfectly accurate feedback). Note that this is significantly below what was observed in the experiment, because the mean-field adaptive algorithm used in the experiment was far from being perfectly accurate.

### Ultimate limit for OPO squeezing and $p = 2$

We have already shown that the lower bound to the MSE for  $p = 2$  scales as  $(\kappa/\mathcal{N})^{2/3}$ . Now we show how to determine the *constant* of the scaling for this model assuming ideal OPO squeezing with  $R_- = 1/R_+$ . We include all terms in Eq. (22) and express our results in terms of  $\mathcal{N}$ . We introduce dimensionless (starred) parameters via  $\mathcal{N} = \kappa\mathcal{N}_*$ ,  $\alpha^2 = \alpha_*^2\kappa\mathcal{N}_*$ ,  $\gamma = \gamma_*\kappa\mathcal{N}_*^{5/6}$ ,  $R_+ = R_*\mathcal{N}_*^{1/3}$ , and  $\omega = \omega_*\kappa\mathcal{N}_*^{2/3}$ , where  $\gamma_* > 0$  and  $R_* > 0$ . Then we obtain in the limit  $\mathcal{N}_* \rightarrow \infty$  (see Appendix F)

$$F^{-1}(0) = \mathcal{N}_*^{-2/3} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_*}{\omega_*^2 + \tilde{F}_*^{(Q)}(\omega_*)}, \quad (25)$$

with

$$\tilde{F}_*^{(Q)}(\omega_*) = \frac{4\gamma_*^2\alpha_*^2}{\gamma_*^2/R_* + \omega_*^2} + \frac{\gamma_*^3\sqrt{R_*}/2}{4\gamma_*^2/R_* + \omega_*^2}. \quad (26)$$

From Eq. (21), the above dimensionless parameters are related by  $\alpha_*^2 = 1 - \gamma_*\sqrt{R_*}/8$ . It is convenient to define  $\tau = \gamma_*\sqrt{R_*}/8$ , so the allowable values for  $\tau$  range from 0

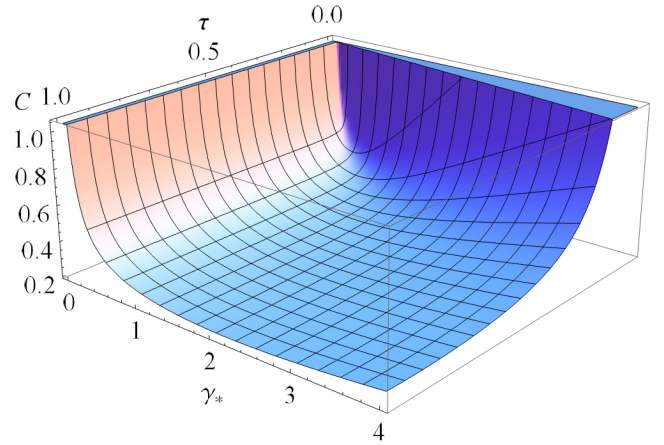


FIG. 1: Plot of  $C$  as a function of  $\gamma_*$  and  $\tau$  (all dimensionless).

to 1. The value of  $C := \mathcal{N}_*^{2/3}F^{-1}(0)$  was calculated for this range of  $\tau$ , and  $\gamma_* \in [0, 4]$ ; the results are given in Fig. 1. It can be seen that  $C$  is smallest for  $\tau = 1$ , which corresponds to a squeezed vacuum, and the minimum is  $C_0 = (587 - 143\sqrt{13})^{1/6}/(4\sqrt{6}) \approx 0.20788$  for  $\gamma_* = 2[2(\sqrt{13} - 3)]^{1/3} \approx 2.1319$ . That is,

$$\langle [\hat{\varphi}(t) - \varphi(t)]^2 \rangle \gtrsim C_0 (\kappa/\mathcal{N})^{2/3}, \quad (27)$$

with  $C_0 \approx 0.20788$ . In this limit of a squeezed vacuum it is only possible to obtain the estimate of the phase modulo  $\pi$ . However, the shallowness of the plot with  $\tau$  shows that one can obtain close to the optimal value for large coherent amplitude, so the phase can be measured modulo  $2\pi$ . The phase tracking simulations in Ref. [4] showed that it is possible to estimate  $\varphi$  modulo  $2\pi$  with  $\langle [\hat{\varphi}(t) - \varphi(t)]^2 \rangle \approx (\kappa/\mathcal{N})^{2/3}$ . Moreover those simulations obtained  $\hat{\varphi}(t)$  by filtering the data prior to  $t$ . By using smoothing [14] of the data before and after  $t$ , one would halve this MSE [1, 5–7].

## CONCLUSION

In summary, we have found a stochastic form of the Heisenberg limit for measurements of a fluctuating phase imposed on a beam with time-invariant statistics. For Wiener fluctuations, the scaling of  $(\kappa/\mathcal{N})^{2/3}$  is tight, in that there is a known adaptive measurement scheme that achieves it. Our bound also reproduces the (tight) Heisenberg scaling of  $(\kappa/\mathcal{N})^{-1}$  for an effectively constant phase. We thus conjecture our general bound to be tight for all power-law phase spectra. We do note, however, that we have assumed a beam with time-symmetric Gaussian statistics, and it is an interesting open question to prove (or disprove) our bound without this assumption.

DWB is funded by an ARC Future Fellowship (FT100100761). HMW and MJWH are supported by

the ARC Centre of Excellence CE110001027.

- 
- [1] H. Yonezawa *et al.*, *Science* **337**, 1514 (2012).  
 [2] D. W. Berry and H. M. Wiseman, *Phys. Rev. A* **65**, 043803 (2002).  
 [3] D. W. Berry and H. M. Wiseman, *Phys. Rev. A* **73**, 063824 (2006).  
 [4] D. W. Berry and H. M. Wiseman, *Phys. Rev. A* **87**, 019901(E) (2013).  
 [5] M. Tsang, J. H. Shapiro, and S. Lloyd, *Phys. Rev. A* **79**, 053843 (2009).  
 [6] T. A. Wheatley *et al.*, *Phys. Rev. Lett.* **104**, 093601 (2010).  
 [7] K. Iwasawa *et al.*, eprint arXiv:1305.0066  
 [8] M. Tsang, *New J. Phys.* **15**, 073005 (2013).  
 [9] M. J. Holland and K. Burnett, *Phys. Rev. Lett.* **71**, 1355 (1993); Z. Y. Ou, *Phys. Rev. A* **55**, 2598 (1997); D. Braun and J. Martin, *Nature Commun.* **2**, 223 (2011).  
 [10] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2010).  
 [11] M. Tsang, H. M. Wiseman, and C. M. Caves, *Phys. Rev. Lett.* **106**, 090401 (2011).  
 [12] M. J. W. Hall, K. Kumar, and M. Reginatto, arXiv:hep-th/0206235 (2002).  
 [13] R. Bhatia, *Positive Definite Matrices* (Princeton University Press, New Jersey, USA, 2007), Sec. 5.5.  
 [14] M. Tsang, *Phys. Rev. Lett.* **102**, 250403 (2009)  
 [15] In Ref. [1] the correlations, denoted  $X_{\pm}(\tau)$ , are symmetrically ordered so that  $X_{\pm}(\tau) = T_{\pm}(\tau) + \delta(\tau)$ .  
 [16] In Ref. [1]  $2\Delta\Omega_0$  is used instead of  $\gamma$ .  
 [17] R. Simon, N. Mukunda, and B. Dutta, *Phys. Rev. A* **49**, 1567 (1994).  
 [18] G. Tóth and D. Petz, *Phys. Rev. A* **87**, 032324 (2013).

## Supplemental Material

### A: SIMPLIFYING THE FISHER INFORMATION FOR GAUSSIAN STATES

Here we show how to obtain Eq. (9) from (5) in the paper. To evaluate averages such as  $\langle X^2(t)X^2(t') \rangle$ , note that the operators for the different times commute, so this is the same as a symmetrically ordered moment, and we can use the Wigner function. This means that the expectation values will satisfy the same rules as for probability distributions.

For Gaussian distributions, one finds that

$$\langle x_1^2 x_2^2 \rangle - \langle x_1^2 \rangle \langle x_2^2 \rangle = 2(\langle x_1 x_2 \rangle^2 - \bar{x}_1^2 \bar{x}_2^2). \quad (\text{S1})$$

As a result,

$$\begin{aligned} \langle Q_{\pm}^2(t)Q_{\pm}^2(t') \rangle &= \langle Q_{\pm}^2 \rangle \langle Q_{\pm}^2 \rangle + 2\langle Q_{\pm}(t)Q_{\pm}(t') \rangle^2 \\ &\quad - 2\langle Q_{\pm} \rangle^2 \langle Q_{\pm} \rangle^2, \end{aligned} \quad (\text{S2})$$

where  $Q_+ := X$  and  $Q_- := Y$ . Note that for expectation values at a single time, we can omit the time-dependence due to stationarity. For  $t \neq t'$ ,  $X(t)$  and  $Y(t')$  commute, so we can again use the Wigner function to take the average, giving

$$\begin{aligned} \langle X^2(t)Y^2(t') \rangle &= \langle X^2 \rangle \langle Y^2 \rangle + 2\langle X(t)Y(t') \rangle^2 \\ &\quad - 2\langle X \rangle^2 \langle Y \rangle^2. \end{aligned} \quad (\text{S3})$$

On the other hand, to include the case  $t = t'$ , we need to explicitly symmetrise. This is not needed for  $\langle X^2(t)X^2(t') \rangle$ , because it is already symmetric, but it is needed for  $\langle X^2(t)Y^2(t') \rangle$ .

Using the commutation relation  $[X(t), Y(t')] = 2i\delta(t - t')$  enables us to evaluate the symmetrised form

$$\begin{aligned} \langle X^2(t)Y^2(t') \rangle_S &= \langle X^2(t)Y^2(t') \rangle \\ &\quad - 4i\delta(t - t')\langle X(t)Y(t') \rangle - 2\delta^2(t - t'). \end{aligned} \quad (\text{S4})$$

The subscript  $S$  indicates the expectation value of the symmetrically ordered operator. We then obtain

$$\begin{aligned} \langle X^2(t)Y^2(t') \rangle &= \langle X^2(t)Y^2(t') \rangle_S + 4i\delta(t - t')\langle X(t)Y(t') \rangle \\ &\quad + 2\delta^2(t - t'). \end{aligned} \quad (\text{S5})$$

Similarly we obtain

$$\begin{aligned} \langle Y^2(t)X^2(t') \rangle &= \langle Y^2(t)X^2(t') \rangle_S - 4i\delta(t - t')\langle Y(t)X(t') \rangle \\ &\quad + 2\delta^2(t - t'). \end{aligned} \quad (\text{S6})$$

Then we get

$$\begin{aligned} \langle X^2(t)Y^2(t') \rangle_S &= \langle X^2 \rangle \langle Y^2 \rangle + 2\langle X(t)Y(t') \rangle_S^2 \\ &\quad - 2\langle X \rangle^2 \langle Y \rangle^2 \\ &= \langle X^2 \rangle \langle Y^2 \rangle + 2\langle X(t)Y(t') \rangle^2 \\ &\quad - 4i\delta(t - t')\langle X(t)Y(t') \rangle - 2\delta^2(t - t') \\ &\quad - 2\langle X \rangle^2 \langle Y \rangle^2, \end{aligned} \quad (\text{S7})$$

and so the delta functions cancel out to give Eq. (S3) even for  $t = t'$ . Similarly

$$\begin{aligned} \langle Y^2(t)X^2(t') \rangle &= \langle Y^2 \rangle \langle X^2 \rangle + 2\langle Y(t)X(t') \rangle^2 \\ &\quad - 2\langle Y \rangle^2 \langle X \rangle^2, \end{aligned} \quad (\text{S8})$$

both for  $t = t'$  and for  $t \neq t'$ .

We can write Eq. (5) in the form

$$\begin{aligned} F^{(Q)}(t - t') &= 4 [\text{Re}\langle b^\dagger(t)b(t)b^\dagger(t')b(t') \rangle \\ &\quad - \langle b^\dagger(t)b(t) \rangle \langle b^\dagger(t')b(t') \rangle] \\ &= \frac{1}{4} \langle [X^2(t) + Y^2(t)][X^2(t') + Y^2(t')] \rangle \\ &\quad - \frac{1}{4} \langle X^2 + Y^2 \rangle^2. \end{aligned} \quad (\text{S9})$$

Using the above results for Gaussian states, Eqs. (S2), (S3), (S8), this becomes

$$\begin{aligned} F^{(Q)}(t, t') &= \frac{1}{2} [\langle X(t)X(t') \rangle^2 + \langle Y(t)Y(t') \rangle^2 \\ &\quad + \langle X(t)Y(t') \rangle^2 + \langle Y(t)X(t') \rangle^2 - (\langle X \rangle^2 + \langle Y \rangle^2)^2]. \end{aligned} \quad (\text{S10})$$

Changing to the normally ordered form we obtain

$$\begin{aligned} F^{(Q)}(t, t') &= 4\mathcal{N}\delta(t - t') - \frac{1}{2} (\langle X \rangle^2 + \langle Y \rangle^2)^2 \\ &\quad + \frac{1}{2} [\langle : X(t)X(t') : \rangle^2 + \langle : Y(t)Y(t') : \rangle^2 \\ &\quad + \langle : X(t)Y(t') : \rangle^2 + \langle : Y(t)X(t') : \rangle^2]. \end{aligned} \quad (\text{S11})$$

Using the functions  $f$  and  $g$ , Eq. (S11) simplifies to Eq. (9) in the paper.

### B: BOUNDING $\tilde{g}(\omega)$ AND $\tilde{f}(\omega)$

Let us define the 2-vector  $Z(t)$ ,  $2 \times 2$  Hermitian matrix function  $M(t, t')$ , and symmetrised correlation function  $S_{AB}(t, t')$  by

$$Z(t) := \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix},$$

$$M(t, t') := \langle Z(t)Z^\dagger(t') + Z(t')Z^\dagger(t) \rangle / 2,$$

$$\begin{aligned} S_{AB}(t, t') &:= \langle A(t)B(t') + B(t')A(t) + A(t')B(t) \\ &\quad + B(t)A(t') \rangle / 4. \end{aligned} \quad (\text{S12})$$

Using the commutation relation  $[X(t), Y(t')] = 2i\delta(t - t')$ , one can write  $M$  in the form

$$M = \begin{pmatrix} S_{XX} & S_{XY} + i\mathbb{1} \\ S_{XY} - i\mathbb{1} & S_{YY} \end{pmatrix}, \quad (\text{S13})$$

where  $\mathbf{1}(t, t') := \delta(t - t')$ . Note that  $M(t, t')$  is positive definite, in the sense that

$$\int dt dt' \xi(t) M(t, t') \xi^\dagger(t') = \langle \aleph_1 \aleph_1^\dagger + \aleph_2 \aleph_2^\dagger \rangle / 2 \geq 0 \quad (\text{S14})$$

for any 2-vector function  $\xi(t)$ , with  $\aleph_1 := \int dt \xi(t) Z(t)$  and  $\aleph_2 := \int dt \xi^\dagger(t) Z(t)$  (this is the continuous analog of the Schrödinger-Robertson multimode uncertainty relation [17]). Stationarity implies further that  $M(t, t')$  can be written as  $M(t - t')$ . Hence, taking Fourier transforms, Eq. (S14) is equivalent to the property

$$\int d\omega \tilde{\xi}(\omega) \tilde{M}(\omega) \tilde{\xi}^\dagger(\omega) \geq 0 \quad (\text{S15})$$

for all 2-vector functions  $\tilde{\xi}(\omega)$ , implying that the  $2 \times 2$  matrix  $\tilde{M}$  is positive; i.e., that

$$\begin{pmatrix} \tilde{S}_{XX} & \tilde{S}_{XY} + i \\ \tilde{S}_{YX} - i & \tilde{S}_{YY} \end{pmatrix} \geq 0 \quad (\text{S16})$$

(a two-dimensional form of Bochner's theorem [13]). This can equivalently be written as the three inequalities  $\tilde{S}_{XX}(\omega) \geq 0$ ,  $\tilde{S}_{YY}(\omega) \geq 0$ , and the spectral uncertainty principle

$$\tilde{S}_{XX}(\omega) \tilde{S}_{YY}(\omega) \geq [\tilde{S}_{XY}(\omega)]^2 + 1. \quad (\text{S17})$$

Note that  $S_{XX}$ ,  $S_{YY}$ , and  $S_{XY}$  are time symmetric by construction, and so have real Fourier transforms.

Let us further define

$$\begin{aligned} g^a(t - t') &:= \langle : X(t) X(t') : \rangle \langle : Y(t) Y(t') : \rangle \\ g^b(t - t') &:= -\frac{1}{2} (\langle : X(t) Y(t') : \rangle^2 + \langle : Y(t) X(t') : \rangle^2) \\ g^c(t - t') &:= \frac{1}{2} (\langle X \rangle^2 + \langle Y \rangle^2)^2, \end{aligned} \quad (\text{S18})$$

so from Eq. (11) of the paper,  $g(t - t') = g^a(t - t') + g^b(t - t') + g^c(t - t')$ . Our first aim is to bound  $\tilde{g}(\omega)$  in terms of  $\mathcal{N}$  using the inequality in Eq. (S17).

The inequality (S17) means that

$$[1 + \tilde{h}_X(\omega)][1 + \tilde{h}_Y(\omega)] \geq [\tilde{h}_{XY}(\omega)]^2 + 1, \quad (\text{S19})$$

where

$$\begin{aligned} h_X(t - t') &:= \langle : X(t) X(t') : \rangle \\ h_Y(t - t') &:= \langle : Y(t) Y(t') : \rangle \\ h_{XY}(t - t') &:= \frac{1}{2} [\langle : X(t) Y(t') + Y(t) X(t') : \rangle]. \end{aligned} \quad (\text{S20})$$

Because  $h_X$ ,  $h_Y$ , and  $h_{XY}$  are time symmetric by construction,  $\tilde{h}_X$ ,  $\tilde{h}_Y$  and  $\tilde{h}_{XY}$  are real. Further, the assumption in the paper that the cross-correlation between the quadratures is time symmetric, i.e., that the symmetrized (or Wigner) cross-correlations satisfy  $\frac{1}{2} \langle X(t) Y(t') + Y(t') X(t) \rangle = \frac{1}{2} \langle X(t') Y(t) + Y(t) X(t') \rangle$ , is equivalent to  $\langle : X(t) Y(t') : \rangle = \langle : X(t') Y(t) : \rangle$ , implying

$$h_{XY}(t - t') = \langle : X(t) Y(t') : \rangle. \quad (\text{S21})$$

Using the above relations, we can write the photon flux as

$$\begin{aligned} \mathcal{N} &= \frac{2}{\pi} \int d\nu [\tilde{h}_X(\nu) + \tilde{h}_Y(\nu)] \\ &= \frac{2}{\pi} \int d\nu [\tilde{h}_X(\omega - \nu) + \tilde{h}_Y(\omega - \nu)]. \end{aligned} \quad (\text{S22})$$

Similarly

$$\begin{aligned} \tilde{g}^a(\omega) &= \frac{1}{2\pi} \int d\nu \tilde{h}_X(\nu) \tilde{h}_Y(\omega - \nu) \\ &= \frac{1}{2\pi} \int d\nu \tilde{h}_X(\omega - \nu) \tilde{h}_Y(\nu), \end{aligned} \quad (\text{S23})$$

and

$$\tilde{g}^b(\omega) = -\frac{1}{2\pi} \int d\nu \tilde{h}_{XY}(\nu) \tilde{h}_{XY}(\omega - \nu). \quad (\text{S24})$$

Equations (S22) and (S23) then give

$$\pi \mathcal{N} + 4\pi \tilde{g}^a(\omega) = \int d\nu \left\{ [1 + \tilde{h}_Y(\nu)][1 + \tilde{h}_X(\omega - \nu)] + [1 + \tilde{h}_Y(\omega - \nu)][1 + \tilde{h}_X(\nu)] - 2 \right\}. \quad (\text{S25})$$

Replacing  $[1 + \tilde{h}_X]$  by the smaller expression  $[\tilde{h}_{XY}^2 + 1]/[1 + \tilde{h}_Y]$  as per Eq. (S19) gives the inequality

$$\pi \mathcal{N} + 4\pi \tilde{g}^a(\omega) \geq \int d\nu \left\{ \frac{[\tilde{h}_Y(\nu) - \tilde{h}_Y(\omega - \nu)]^2}{[1 + \tilde{h}_Y(\nu)][1 + \tilde{h}_Y(\omega - \nu)]} + [\tilde{h}_{XY}(\nu)]^2 \frac{1 + \tilde{h}_Y(\omega - \nu)}{1 + \tilde{h}_Y(\nu)} + [\tilde{h}_{XY}(\omega - \nu)]^2 \frac{1 + \tilde{h}_Y(\nu)}{1 + \tilde{h}_Y(\omega - \nu)} \right\}. \quad (\text{S26})$$

Each of the three terms here is positive. Dropping the

first term and using the inequality  $a + b \geq 2\sqrt{ab}$ , gives

$$\pi \mathcal{N} + 4\pi \tilde{g}^a(\omega) \geq 2 \int d\nu |\tilde{h}_{XY}(\nu) \tilde{h}_{XY}(\omega - \nu)|. \quad (\text{S27})$$

Because the absolute value cannot exceed the value, it follows that  $\mathcal{N} + 4\tilde{g}^a(\omega) + 4\tilde{g}^b(\omega) \geq 0$ . Moreover,  $g^c(t - t')$  is a positive constant by stationarity, and therefore  $\tilde{g}^c(\omega) = 2\pi g^c(0)\delta(\omega) \geq 0$ . This means that  $\mathcal{N} + 4\tilde{g}(\omega) \geq 0$ .

Turning now to  $\tilde{f}(\omega)$ , from Eq. (10) we can similarly write this as

$$\tilde{f}(\omega) = \int d\nu \left[ \tilde{h}_X(\nu) + \tilde{h}_Y(\nu) \right] \left[ \tilde{h}_X(\omega - \nu) + \tilde{h}_Y(\omega - \nu) \right]. \quad (\text{S28})$$

Using  $\tilde{h}_X(\omega) + \tilde{h}_Y(\omega) = \tilde{h}_X(\omega) + [1 + \tilde{h}_Y(\omega)] - 1$  and replacing  $1 + \tilde{h}_Y(\omega)$  by the smaller term  $1/[1 + \tilde{h}_X(\omega)]$ , as per the weaker version of Eq. (S19) with 1 on the right-hand-side, and simplifying yields

$$\tilde{h}_X(\omega) + \tilde{h}_Y(\omega) \geq \frac{[\tilde{h}_X(\omega)]^2}{1 + \tilde{h}_X(\omega)} \geq 0. \quad (\text{S29})$$

As  $\tilde{f}(\omega)$  is thus an integral of a product of two non-negative quantities, we immediately get  $\tilde{f}(\omega) \geq 0$ .

### C: DETAILS OF THE INEQUALITIES IN EQ. (15)

The inequality in the second line of Eq. (15) of the paper is because the integral over the interval  $[0, 2\mathcal{I}/\mu]$  cannot be more than half that over the full range, as we now prove. The integral over the full range  $[0, \infty)$  is given by

$$\begin{aligned} & \frac{1}{\pi} \int_0^\infty \frac{\kappa^{p-1} d\omega}{\omega^p + \kappa^{p-1} (\zeta\mathcal{N} + \mu)} \\ &= \frac{\kappa^{p-1}}{\pi [\kappa^{p-1} (\zeta\mathcal{N} + \mu)]^{1-1/p} \text{sinc}(\pi/p)} \\ &\geq \frac{\kappa^{p-1}}{\pi [\kappa^{p-1} (\zeta\mathcal{N} + \mu)]^{1-1/p}}. \end{aligned} \quad (\text{S30})$$

In addition, if  $(\mathcal{I}/\mu)^p \leq \kappa^{p-1} (\zeta\mathcal{N} + \mu)$ , then

$$\begin{aligned} & \frac{1}{\pi} \int_0^{\mathcal{I}/2\mu} \frac{\kappa^{p-1} d\omega}{\omega^p + \kappa^{p-1} (\zeta\mathcal{N} + \mu)} \\ &\leq \frac{1}{\pi} \int_0^{[\kappa^{p-1} (\zeta\mathcal{N} + \mu)]^{1/p/2}} \frac{\kappa^{p-1} d\omega}{\kappa^{p-1} (\zeta\mathcal{N} + \mu)} \\ &= \frac{\kappa^{p-1}}{2\pi [\kappa^{p-1} (\zeta\mathcal{N} + \mu)]^{1-1/p}}, \end{aligned} \quad (\text{S31})$$

Comparing this to Eq. (S30) gives the required result.

### D: MORE GENERAL TYPES OF PHASE FLUCTUATIONS

We allow for phase fluctuations satisfying the condition

$$\exists \beta : \forall |\omega| > \kappa\beta, \tilde{F}^{(C)}(\omega) \leq |\omega|^p / \kappa^{p-1}. \quad (\text{S32})$$

Note that we continue to assume that the correlations are stationary, as per condition 1. For *Gaussian* phase statistics,  $\tilde{F}^{(C)}(\omega) = \tilde{\Sigma}^{-1}(\omega)$ , although more generally  $\tilde{F}^{(C)}(\omega) \geq \tilde{\Sigma}^{-1}(\omega)$  (since  $F^{(C)}(t, t') - \Sigma^{-1}(t, t')$  is a positive definite function [12]). Our condition means that, for large  $\omega$ ,  $\tilde{\Sigma}(\omega)$  is greater than or equal to a function scaling as  $\kappa^{p-1}/|\omega|^p$ . We could also give the condition directly in terms of  $\tilde{\Sigma}(\omega)$  and it would be equivalent for Gaussian phase statistics, but this form of the condition also allows for non-Gaussian phase statistics in Eq. (7). A particular example, for  $p = 2$  (and Gaussian statistics), is the case of phase fluctuations modelled by Ornstein-Uhlenbeck noise, where  $\tilde{\Sigma}(\omega) = \kappa/(\lambda^2 + \omega^2)$ .

Selecting a suitable value of  $\beta$ , then for  $|\omega| > \kappa\beta$ ,

$$\tilde{F}^{-1}(\omega) \geq \frac{1}{|\omega|^p / \kappa^{p-1} + \tilde{F}^{(Q)}(\omega)}. \quad (\text{S33})$$

Following the same derivation as in the body of the paper, we obtain

$$F^{-1}(0) \geq \frac{1}{\pi} \int_{\mathcal{I}/2\mu + \kappa\beta}^\infty \frac{\kappa^{p-1} d\omega}{\omega^p + \kappa^{p-1} (\zeta\mathcal{N} + \mu)}. \quad (\text{S34})$$

Here, we now omit the smallest range of integration consistent with  $|\omega| > \kappa\beta$  (as Eq. (S33) requires  $|\omega| > \kappa\beta$ ), as well as the range  $|\omega| \leq \kappa\beta$ . As in the body of the paper, omitting part of the range of the integral can only further reduce the value of the integral.

As we take  $\mu = \Theta(\mathcal{N}(\mathcal{N}/\kappa)^{(p-1)/(p+1)})$  and consider large  $\mathcal{N}/\kappa$ , we obtain  $\mathcal{I}/2\mu > \kappa\beta$ , so we again obtain Eq. (15) from the paper. Therefore the more general type of phase fluctuations we have allowed do not alter the scaling.

### E: THE FISHER INFORMATION FOR MIXED STATES

A mixed phase-dependent state can be written in the form

$$\rho(\varphi) = \sum_\ell p_\ell \rho_\ell(\varphi). \quad (\text{S35})$$

Consider the Fisher information for estimation of the unknown phase shift  $\varphi$  via the optimal measurement with positive operator-valued measure (POVM)  $\{M_k\}$ . We use  $q_k$  for the probability of measurement result  $k$  on  $\rho(\varphi)$ . To upper bound the Fisher information obtainable by this POVM, we can always consider a more informative measurement described by POVM  $\{|\ell\rangle\langle\ell| \otimes M_k\}$  on the expanded state  $\rho(\varphi) = \sum_\ell p_\ell |\ell\rangle\langle\ell| \otimes \rho_\ell(\varphi)$ . We use  $p_{k\ell}$  to indicate the probability of measurement result  $(k, \ell)$  with this expanded POVM, and  $p_{k|\ell} = p_{k\ell}/p_\ell$ . Because the expanded POVM cannot result in less infor-



mation, we must have [18]

$$\begin{aligned}
F^{(Q)}(\rho(\varphi)) &= \sum_k q_k \left[ \frac{d}{d\varphi} \log(q_k) \right]^2 \\
&\leq \sum_{k\ell} p_{k\ell} \left[ \frac{d}{d\varphi} \log(p_{k\ell}) \right]^2 \\
&= \sum_\ell p_\ell \sum_k p_{k|\ell} \left[ \frac{d}{d\varphi} \log(p_{k|\ell}) \right]^2 \\
&\leq \sum_\ell p_\ell F_Q(\rho_\ell(\varphi)). \tag{S36}
\end{aligned}$$

As a result, the Fisher information is no greater than that for the individual states in the mixture.

When considering a mixture of phase-squeezed states, we have no  $XY$  correlations, and therefore require  $S_{XX}S_{YY} \geq \mathbb{1}$ . We are considering the case that the  $Y$  quadrature is squeezed, and there is antisqueezing in the  $X$  quadrature. For mixed squeezed states, we can model the  $X$  fluctuations as a sum of classical and quantum fluctuations. That is, we define  $S_{XX}^Q := \mathbb{1}/S_{YY}$  and  $S_{XX}^C := S_{XX} - S_{XX}^Q$ , so  $S_{XX} = S_{XX}^Q + S_{XX}^C$ . The inequality  $S_{XX}S_{YY} \geq \mathbb{1}$  means that  $S_{XX}^C$  is positive semidefinite, and therefore represents valid classical fluctuations.

Moreover, it is easily shown that the correlations  $S_{XX}^Q$  and  $S_{YY}$  correspond to a pure squeezed state. For this pure squeezed state, we take  $R_+^Q = 1/R_-$ , and adjust the values of  $x$  and  $\gamma$  to be consistent. That is,  $x^Q$  and  $\gamma^Q$  such that  $x^Q$  is determined from  $R_+^Q$  and  $R_-$ , and  $\gamma^Q$  satisfies  $\gamma^Q(1+x^Q) = \gamma(1+x)$ .

## F: DERIVATION OF EQ. (25)

Substituting the equations with the starred quantities into the equation for  $F^{-1}(0)$ , one obtains

$$F^{-1}(0) = \mathcal{N}_\star^{-2/3} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_\star}{\omega_\star^2 + [\lambda/\kappa + \tilde{F}^{(Q)}(\omega)]/\kappa\mathcal{N}_\star^{4/3}}. \tag{S37}$$

We use Eq. (22) of the paper, and take  $R_- = 1/R_+$  and  $x = (\sqrt{R_+} - 1)/(\sqrt{R_+} + 1)$ . In the limit of large  $R_+$  we have  $x = 1 - 2R_+^{-1/2} + O(R_+^{-1})$ . Keeping only the leading order term, we have  $1 - x \approx 2R_+^{-1/2}$ . Similarly, keeping only leading order terms we have  $R_+ - 1 \approx R_+$ ,  $R_- - 1 \approx -1$ , and  $1 + x \approx 2$ . Then Eq. (22) simplifies to

$$\begin{aligned}
\tilde{F}^{(Q)}(\omega) &\approx 4\mathcal{N} + 4\alpha^2 R_+ \frac{(4/R_+)\gamma^2}{(4/R_+)\gamma^2 + 4\omega^2} \\
&+ \frac{\gamma^3}{16} \left[ \frac{8\sqrt{R_+}}{(4/R_+)\gamma^2 + \omega^2} + \frac{8}{4\gamma^2 + \omega^2} \right]. \tag{S38}
\end{aligned}$$

For large  $R_+$  the second term in the square brackets can be omitted. Substituting the starred quantities then gives

$$\begin{aligned}
\tilde{F}^{(Q)}(\omega) &\approx 4\kappa\mathcal{N}_\star + 4\alpha_\star^2\kappa\mathcal{N}_\star^{4/3} \frac{\gamma_\star^2}{\gamma_\star^2/R_\star + \omega_\star^2} \\
&+ \gamma_\star^3\kappa\mathcal{N}_\star^{4/3} \frac{\sqrt{R_\star}/2}{4\gamma_\star^2/R_\star + \omega_\star^2}. \tag{S39}
\end{aligned}$$

Now we take  $\tilde{F}_\star^{(Q)}(\omega_\star) := \tilde{F}^{(Q)}(\omega)/\kappa\mathcal{N}_\star^{4/3}$ , so that

$$\tilde{F}^{(Q)}(\omega) \approx 4\mathcal{N}_\star^{-1/3} + \frac{4\gamma_\star^2\alpha_\star^2}{\gamma_\star^2/R_\star + \omega_\star^2} + \frac{\gamma_\star^3\sqrt{R_\star}/2}{4\gamma_\star^2/R_\star + \omega_\star^2}. \tag{S40}$$

The first term is higher order, and omitting it gives Eq. (26) of the paper. Omitting the term in  $\lambda$  and using  $\tilde{F}_\star^{(Q)}(\omega_\star)$  in Eq. (S37) then gives Eq. (25) of the paper.