# Quantum-Enhanced Optical-Phase Tracking

## Author

Yonezawa, Hidehiro, Nakane, Daisuke, Wheatley, Trevor A, Iwasawa, Kohjiro, Takeda, Shuntaro, Arao, Hajime, Ohki, Kentaro, Tsumura, Koji, Berry, Dominic W, Ralph, Timothy C, Wiseman, Howard M, Huntington, Elanor H, Furusawa, Akira

## Published

## Journal Title

## DOI

## Copyright Statement

## Downloaded from

## Griffith Research Online

# Quantum-enhanced optical phase tracking

Hidehiro Yonezawa,[1] Daisuke Nakane,[1] Trevor A. Wheatley,[1,2,3]
Kohjiro Iwasawa,[1] Shuntaro Takeda,[1] Hajime Arao,[1] Kentaro Ohki,[4]
Koji Tsumura,[5] Dominic W. Berry,[6,7] Timothy C. Ralph,[2,8]
Howard M. Wiseman,[2,9*] Elanor H. Huntington,[2,3] Akira Furusawa[1*]

[1]Department of Applied Physics, School of Engineering,
The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[2]Centre for Quantum Computation and Communication Technology, Australian Research Council
[3]School of Engineering and Information Technology, University College,
The University of New South Wales, Canberra ACT 2600, Australia
[4]Department of Applied Mathematics and Physics, Graduate School of Informatics,
Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan
[5]Department of Information Physics and Computing,
The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[6]Institute for Quantum Computing, University of Waterloo, Ontario N2L 3G1, Canada
[7]Department of Physics and Astronomy, Macquarie University, NSW 2109, Australia
[8]School of Mathematics and Physics, University of Queensland, Brisbane QLD 4072, Australia
[9]Centre for Quantum Dynamics, Griffith University, Brisbane QLD 4111, Australia

*To whom correspondence should be addressed. E-mail:
h.wiseman@griffith.edu.au (H.W.); akiraf@ap.t.u-tokyo.ac.jp (A.F.)

**Tracking a randomly varying optical phase is a key task in metrology, with applications in optical communication. The best precision for optical phase tracking has till now been limited by the quantum vacuum fluctuations of coherent light. Here we surpass this coherent-state limit by using a continuous-wave beam in a phase-squeezed quantum state. Unlike in previous squeezing-**

1

**enhanced metrology, restricted to phases with very small variation, the best tracking precision (for a fixed light intensity) is achieved for a finite degree of squeezing, due to Heisenberg's uncertainty principle. By optimizing the squeezing we track the phase with a mean square error 15±4% below the coherent-state limit.**

There are many tasks where precise optical phase estimation is critical, including communication (*1, 2*), and metrology (*3*). Quantum mechanics imposes a fundamental bound on precision (*4–6*), and this already limits gravitational wave detection (*7–9*) and can guarantee security in quantum cryptography (*10*). The quantum limits are determined by optimizing (subject to constraints) the input quantum state, the quantum measurement, and the data processing. Much effort has been devoted to approaching the fundamental quantum limits (*3, 5, 6*).

Phase estimation can be divided into two kinds (*11*): phase sensing, where the phase is known to always lie within some small interval (e.g. (*12*)), and general phase estimation, where it is not so constrained (e.g. (*11*)). In the former case, when (as in most practical situations) the field has a large coherent amplitude, the problem can be linearized in terms of the phase rotation (*7*), which greatly simplifies the task of optimizing the input state and the measurement. By contrast, in the case of unconstrained phase estimation the problem cannot be linearized, and as a consequence the optimization problem is considerably harder (*11, 13–22*). While a quantum enhancement of phase sensing using nonclassical states of light has recently been demonstrated (*8, 9*), this has been done for general phase estimation only with post-selected results (*11*).

We present a demonstration of unconstrained phase estimation with a quantum enhancement using nonclassical (squeezed) states. We use homodyne detection, with no post-selection of data, and no compensation for losses or detector inefficiency in the system. Moreover, the problem of a stochastically varying phase is addressed (*8, 9, 19–22*), as is highly relevant for

physical metrology and communication, rather than a time-invariant (but initially unknown) phase (*11–13, 15–18*). To perform optimal estimation we have implemented optical phase tracking — a phase-lock loop which strives to maintain the maximum measurement sensitivity for a widely varying phase. The quantum noise in the photocurrent prevents the maximum sensitivity from being perfectly maintained, which is why we observe an optimal degree of squeezing.

Our experiment (Fig. 1A) uses a continuous-wave optical phase-squeezed beam. The phase of the beam is modulated with the signal $\varphi(t)$, the waveform to be estimated (*22–24*). As in Refs. (*21, 22*), we use a stochastic waveform defined by

$$\varphi(t) = \sqrt{\kappa} \int_{-\infty}^{t} e^{-\lambda(t-s)} dV(s). \tag{1}$$

Here $dV(s)$ is a classically generated Wiener process (*25*) (white noise), $\lambda^{-1}$ is the correlation time of $\varphi(t)$, and $\kappa$ determines the magnitude of the phase variation, which is of order unity. This $\varphi(t)$ is a continuous-time random walk with a tendency to return to the mean phase of zero, a kind of noisy relaxation process that occurs in many physical situations (*25*).

The phase-modulated beam is measured by homodyne detection, using a local oscillator (LO), yielding a noisy current $I(t)$. The LO phase $\Phi(t)$ is feedback-controlled to be $\Phi(t) \approx \pi/2 + \varphi(t)$, as this is the most sensitive operating point for sensing changes in $\varphi(t)$ (Fig. 1B). Because $\varphi(t)$ is unknown, the best strategy is adaptive metrology (*11, 13, 15–21*), in which feedback control is used to set $\Phi(t) = \varphi_f(t) + \pi/2$, where $\varphi_f(t)$ is a filtered estimate of $\varphi(t)$ — that is, an estimate based on $I(s)$ for all $s < t$. This gives a normalized homodyne output current $I(t)$ of (*19, 26*),

$$I(t)dt \simeq 2\,|\alpha|\,[\varphi(t) - \varphi_f(t)]\,dt + \sqrt{\bar{R}_{\text{sq}}}\,dW(t),$$

$$\bar{R}_{\text{sq}} = \sigma_f^2 e^{2r_p} + (1 - \sigma_f^2)e^{-2r_m}. \tag{2}$$

Here $|\alpha|$ is the amplitude of the input phase-squeezed beam, and $dW(t)$ is another Wiener process (*25*), arising from the squeezed vacuum fluctuations. The magnitude $\bar{R}_{\text{sq}}$ of these quantum

3

fluctuations is determined by the degree of squeezing ($r_m \geq 0$) and anti-squeezing ($r_p \geq r_m$), and by $\sigma_f^2 = \langle [\varphi(t) - \varphi_f(t)]^2 \rangle$. Note that several approximations — justified in Ref. (26) — have been made to derive Eq. 2, most notably, a second-order expansion for $I(t)dt$ in the small variable $[\varphi(t) - \varphi_f(t)]$.

For optimal feedback control, the Kalman filter is used for $\varphi_f(t)$ (22), which is the causal (i.e. real time) estimator with the lowest mean square error (MSE). The Kalman filter is the optimal filter for estimating $\varphi(t)$ of the form of Eq. 1 when using a coherent beam (22), and the calculation generalizes to our squeezed case (26). Though the filtered estimate $\varphi_f(t)$ is a good estimate of the signal phase $\varphi(t)$, to obtain the best estimate we apply the acausal technique of smoothing (21, 22, 24). After storing data over a certain period of time, a precise estimate of $\varphi_s(t)$ is obtained at a time $t$ in the middle of that period by using observations both before and after $t$. The MSE of the smoothed estimate $\sigma_s^2 = \langle [\varphi(t) - \varphi_s(t)]^2 \rangle$ is given as (22, 26),

$$\sigma_s^2 = \kappa / \left( 2\sqrt{4\kappa |\alpha|^2 / \bar{R}_{\text{sq}} + \lambda^2} \right). \tag{3}$$

Recall that $\bar{R}_{\text{sq}}$ (Eq. 2) is a function of $\sigma_f^2$, so the above expression for $\sigma_s^2$ is still implicit. The full solutions are given in (26), but in the parameter regime of our experiment, $\sigma_s^2$ is approximately proportional to $\sqrt{\bar{R}_{\text{sq}}}$. That is, by using a nonclassical beam with effective squeezing $\bar{R}_{\text{sq}} < 1$ we expect to be able to overcome the coherent-state limit (CSL) by a factor of $\sqrt{\bar{R}_{\text{sq}}}$.

Our experiment (see Fig. 1C) uses an 860 nm continuous-wave Titanium Sapphire laser. The phase-squeezed beam is added by an optical parametric oscillator (OPO). The OPO is driven below threshold by a 430 nm pump beam, generated by a second-harmonic-generation cavity. We obtain up to $-4$ dB of phase squeezing. The signal $\varphi(t)$ is produced by a digital random signal generator and a low-pass filter with a cutoff frequency of $\lambda/2\pi$. This is imposed upon the phase-squeezed beam using an electro-optic modulator. Homodyne detection is performed on this phase-modulated beam with an overall efficiency of $\eta = 0.85$. The homodyne current

4

goes to the optimal feedback filter (another low-pass filter with a cutoff frequency $\lambda/2\pi$ (26)). Its output, $\varphi_f(t)$, is then shifted by $\pi/2$ and applied on the phase of the LO beam with another electro-optic modulator.

We record $\varphi(t)$, $I(t)$ and $\varphi_f(t)$ by an oscilloscope with a sampling rate of 100 MHz. Figure 2 shows a typical segment of the recorded signals, plus the smoothed estimate $\varphi_s(t)$. The parameters here are $\kappa = (1.9 \pm 0.1) \times 10^4$ rad/s, $\lambda = (5.9 \pm 0.5) \times 10^4$ rad/s, $|\alpha|^2 = (1.00 \pm 0.06) \times 10^6$ s$^{-1}$, squeezing $-3.1\pm0.1$ dB ($r_m = 0.36 \pm 0.01$) and anti-squeezing $5.1\pm0.1$ dB ($r_p = 0.59 \pm 0.01$), from a pump beam power of 80 mW. Note that $\kappa$ and $\lambda$ are fixed through this paper. The current $I(t)$ has zero mean because the feedback loop is designed to operate the homodyne measurement at this point of maximum sensitivity (Fig. 1B). While the filtered estimate $\varphi_f(t)$ has a visible delay due to its causal nature, the smoothed estimate $\varphi_s(t)$ does not, and the signal phase $\varphi(t)$ is reliably tracked.

To investigate the squeezing-enhancement, we perform phase tracking with a fixed square amplitude $|\alpha|^2 = (1.00 \pm 0.06) \times 10^6$ s$^{-1}$ but with varying squeezing levels arising from OPO pump beam powers of 0, 30, 80 and 180 mW. Independently of the phase estimation, squeezing and anti-squeezing levels were measured for each pump beam power (26). The red crosses in Fig. 3 show the MSEs of the smoothed estimates $\sigma_s^2$ as a function of the squeezing level. The MSE was calculated from 2 ms of data ($2 \times 10^5$ samples). Repeating this 15 times gave the average MSE and its standard deviation.

Figure 3 shows three key results. First, the squeezing-enhancement is verified: the MSEs are reduced below the CSL (i) by using phase-squeezed beams. Second, the experimental results are in good agreement with the prediction (ii), and in disagreement with the theory curve (iii) which is based on approximating the homodyne output current $I(t)$ to only first order in $[\varphi(t) - \varphi_f(t)]$ so that $\bar{R}_{\text{sq}} = e^{-2r_m}$. Third, at the higher squeezing level the MSE is saturated, indicating the existence of an optimal squeezing level. Even in the theoretical curve (iv) for pure squeezed

5

beams and zero loss, the MSE has a minimum. This curve corresponds to the fundamental limit imposed by Heisenberg's uncertainty principle for the phase and amplitude quadatures, namely $e^{-2r_p} \times e^{2r_m} \geq 1$. Although more squeezing decreases the $e^{-2r_m}$ term in $\bar{R}_{sq}$, it increases the $e^{2r_p}$ term due to the tracking being imperfect, which itself is a consequence of the noise in the photocurrent (2). This defines (self-consistently) the optimal degree of squeezing, which depends upon the parameters $|\alpha|$, $\kappa$ and $\lambda$ (26).

Experimentally, we varied the amplitude $|\alpha|$, while fixing the pump beam power to 80 mW, giving squeezing and anti-squeezing levels of $-3.2\pm0.2$ dB and $4.9\pm0.3$ dB, respectively. Theoretically, the optimal squeezing level increases with $|\alpha|$, and so too does the squeezing-enhancement, without limit. However, for our experimental conditions ($10^6 s^{-1} \leq |\alpha|^2 < 10^7 s^{-1}$) the effect of keeping the squeezing fixed is minor (less than $3\%$ difference to $\sigma_s^2$). Figure 4 shows the dependence of the MSE $\sigma_s^2$ on $|\alpha|$. The theoretical curves show good agreement with experiments. Over the whole amplitude range, the estimates with the squeezed beams surpass what is possible with coherent states, with $\sigma_s^2$, averaged over the four different amplitudes, being $(15 \pm 4)\%$ below the CSL. The conclusion is essentially unaltered if one calculates the CSL not in terms of $|\alpha|^2$, but in terms of the effective photon flux $\mathcal{N}_{eff}$, which equals $|\alpha|^2$ plus the extra photons due to the squeezed vacuum fluctuations in the relevant spectral range (26).

We have tracked the phase of a squeezed optical field that varies stochastically in time over a significant angular range. Our use of Kalman filtering in real-time adaptive measurements of nonclassical systems could be applied also in solid-state and nanomechanical devices. Optimizing both the degree of squeezing and its bandwidth according to the experimental conditions would allow a completely rigorous treatment of photon flux. Lower losses and more squeezing would then enable a dramatic improvement to a precision that scales differently with photon flux, with $\sigma^2 \propto \mathcal{N}^{-5/8}$ (20) as opposed to the $\sigma^2 \propto \mathcal{N}^{-1/2}$ in the current setup.

# References and Notes

1. R. Slavik *et al.*, All-optical phase and amplitude regenerator for next-generation telecommunications systems. *Nature Photon.* **4**, 690-695 (2010).

2. Jian Chen, Jonathan L. Habif, Zachary Dutton, Richard Lazarus, Saikat Guha, Optical codeword demodulation with error rates below the standard quantum limit using a conditional nulling receiver. *Nature Photon.* **6**, 374-379 (2012).

3. V. Giovannetti, S. Lloyd, L. Maccone, Advances in quantum metrology. *Nature Photon.* **5**, 222-229 (2011).

4. C. W. Helstrom, Quantum Detection and Estimation Theory (Academic Press, New York, 1976).

5. V. Giovannetti, S. Lloyd, L. Maccone, Quantum-Enhanced Measurements: Beating the Standard Quantum Limit. *Science* **306**, 1330-1336 (2004).

6. H. M. Wiseman, G. J. Milburn, Quantum Measurement and Control (Cambridge University Press, Cambridge, England, 2010).

7. C. M. Caves, Quantum-mechanical noise in an interferometer. *Phys. Rev. D* **23**, 1693-1708 (1981).

8. K. Goda *et al.*, A quantum-enhanced prototype gravitational-wave detector. *Nature Phys.* **4**, 472-476 (2008).

9. The LIGO Scientific Collaboration, A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nature Phys.* **7**, 962-965 (2011).

10. K. Inoue, E. Waks, Y. Yamamoto, Differential Phase Shift Quantum Key Distribution. *Phys. Rev. Lett.* **89**, 037902 (2002).

11. G. Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, G. J. Pryde, Entanglement-enhanced measurement of a completely unknown optical phase. *Nature Photon.* **5**, 43-47 (2011).

12. T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, S. Takeuchi, Beating the Standard Quantum Limit with Four-Entangled Photons. *Science* **316**, 726-729 (2007).

13. H. M. Wiseman, Adaptive phase measurements of optical modes: going beyond the marginal Q distribution. *Phys. Rev. Lett.* **75**, 4587-4590 (1995).

14. H. M. Wiseman, R. B. Killip, Adaptive single-shot phase measurements: The full quantum theory. *Phys. Rev. A* **57**, 2169-2185 (1998).

15. D.W. Berry, H. M. Wiseman, Optimal States and Almost Optimal Adaptive Measurements for Quantum Interferometry. *Phys. Rev. Lett.* **85**, 5098-5101 (2000).

16. M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, H. Mabuchi, Adaptive homodyne measurement of optical phase. *Phys. Rev. Lett.* **89**, 133602 (2002).

17. B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation. *Nature* **450**, 393-396 (2007).

18. A. Hentschel, B. C. Sanders, Machine Learning for Precise Quantum Measurement. *Phys. Rev. Lett.* **104**, 063603 (2010).

19. D. W. Berry, H. M. Wiseman, Adaptive quantum measurements of a continuous varying phase. *Phys. Rev. A* **65**, 043803 (2002).

20. D. W. Berry, H. M. Wiseman, Adaptive phase measurements for narrowband squeezed beams. *Phys. Rev. A* **73**, 063824 (2006).

21. T. A. Wheatley *et al.*, Adaptive optical phase estimation using time-symmetric quantum smoothing. *Phys. Rev. Lett.* **104**, 093601 (2010).

22. M. Tsang, J. H. Shapiro, S. Lloyd, Quantum theory of optical temporal phase and instantaneous frequency. II. Continuous-time limit and state-variable approach to phase-locked loop design. *Phys. Rev. A* **79**, 053843 (2009).

23. M. Tsang, H. M. Wiseman, C. M. Caves, Fundamental Quantum Limit to Waveform Estimation. *Phys. Rev. Lett.* **106**, 090401 (2011).

24. M. Tsang, Time-symmetric quantum theory of smoothing. *Phys. Rev. Lett.* **102**, 250403 (2009).

25. C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1985).

26. See supplementary materials.

27. Y. Takeno, M. Yukawa, H. Yonezawa, A. Furusawa, Observation of -9 dB quadrature squeezing with improvement of phase stability in homodyne measurement. *Opt. Express* **15**, 4321-4327 (2007).

28. J. Appel, D. Hoffman, E. Figueroa, A. I. Lvovsky, Electronic noise in optical homodyne tomography. *Phys. Rev. A* **75**, 035802 (2007).

29. U. Leonhardt, Essential Quantum Optics (Cambridge, 2010).

30. C. W. Gardiner, P. Zoller, Quantum Noise (Springer, 2000).

**Fig. 1** Experimental setup. (**A**) Schematic of the experiment. (**B**) Homodyne output current versus relative phase between LO and phase-squeezed beam, and phasor diagram for a slightly non-optimal relative phase (as will occur in the phase tracking problem). (**C**) Detail of the experimental setup. The abbreviations are TiS: Titanium Sapphire, CW: continuous-wave, PM: phase modulator, SHG: second-harmonic-generation, OPO: optical parametric oscillator.

**Fig. 2** Time domain results of phase estimate. (**A**) Signal phase to be estimated $\varphi(t)$. (**B**) Homodyne output current $I(t)$. (**C**) Filtered estimate $\varphi_f(t)$. (**D**) Smoothed estimate $\varphi_s(t)$.

**Fig. 3** Smoothed MSE $\sigma_s^2$ versus squeezing level. Red crosses represent experimental data. Trace (i) is the coherent-state limit which is reachable with a coherent beam only if we have unit-detection efficiency $\eta = 1$. Trace (ii) is the theoretical curve from Eq. 3. Trace (iii) is the theoretical curve based on a approximating the homodyne output current $I(t)$ to only first order in $[\varphi(t) - \varphi_f(t)]$ so that $\bar{R}_{\rm sq} = e^{-2r_m}$. Trace (iv) is the theoretical curve from Eq. 3 for pure squeezed beams (i.e. without loss).

**Fig. 4** Dependence of the smoothed MSE $\sigma_s^2$ on the amplitude squared $|\alpha|^2$. Blue and red crosses are experimental data for coherent and squeezed beams, respectively. Trace (i) is the coherent-state limit. Trace (ii) is the theoretical curve for coherent beams with the experimental setup (i.e. including inefficiency). Trace (iii) is the theoretical curve for squeezed beams, including inefficiency. Trace (iv) is the theoretical curve for pure squeezed beams and 100% efficiency, with the squeezing level optimized for each $|\alpha|^2$.
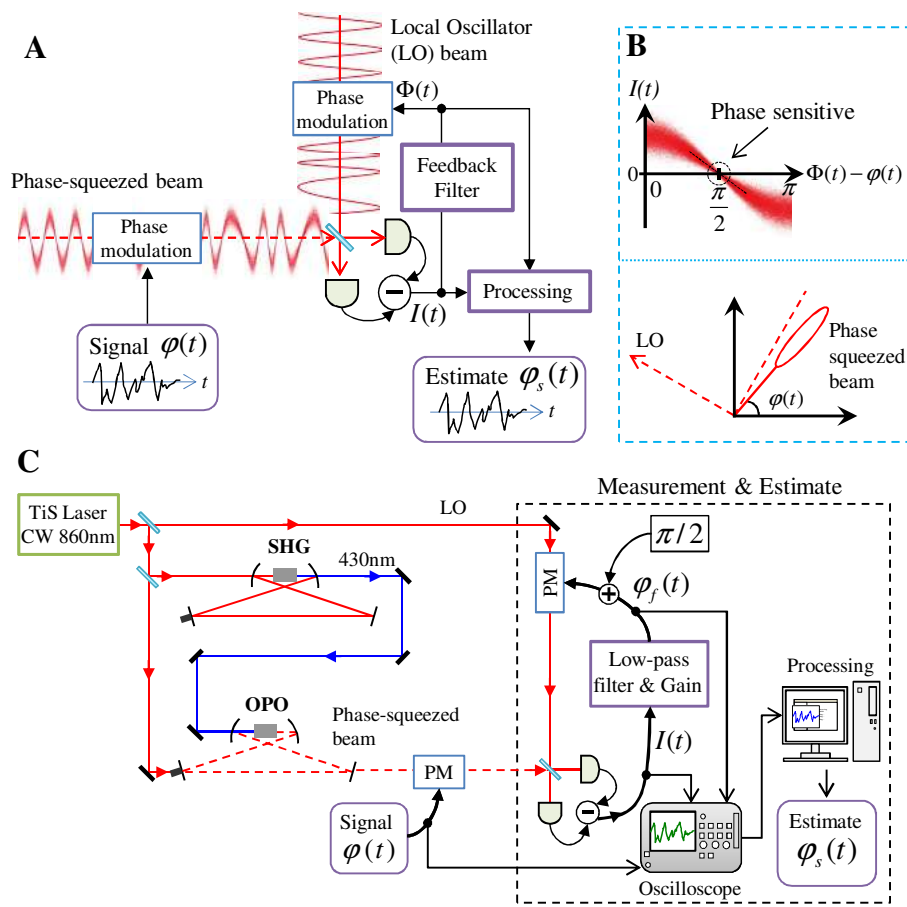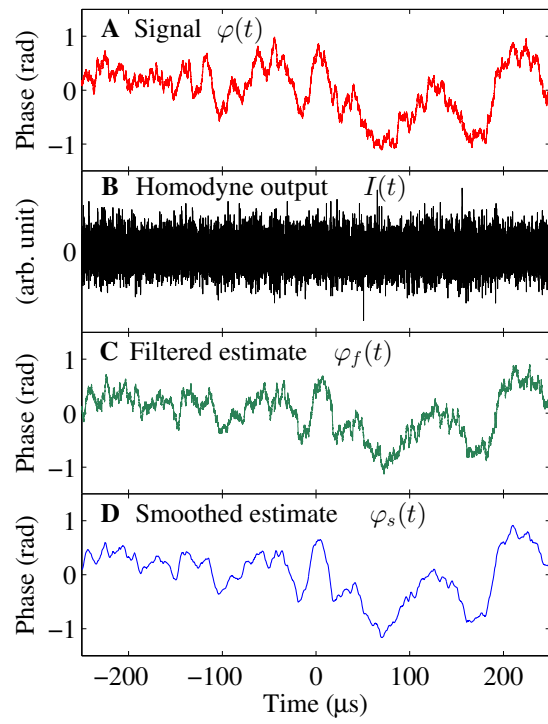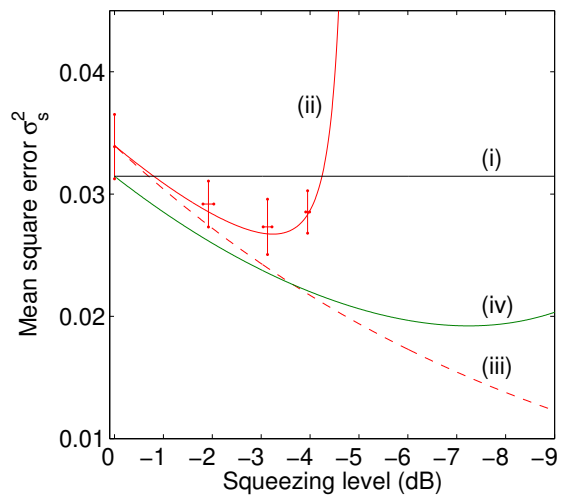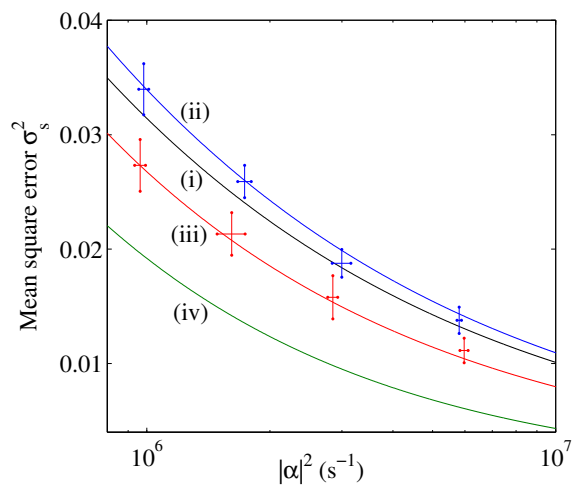
Figure 1:

Figure 2:



Figure 3:

13

Figure 4:

# Supplementary Materials for
# Quantum-enhanced optical phase tracking

## S1  Introduction

Specific details of a more technical nature omitted in the main text are briefly discussed here. Additionally, the following sections cover a more technical discussion the experiment and details of the mathematics supporting the arguments in the main text.

Optical phase estimation has been extensively investigated because of its great importance in both science and engineering. As stated in the main text, phase estimation can be divided into phase sensing, where the phase is known to always lie within some small interval, and unconstrained or general phase estimation, where it is not so constrained (e.g. (*11*)). In phase sensing, if (as is typically the case) the field has a large coherent amplitude, the problem can be linearized in terms of the unknown phase-rotation (*7*). This means that the unknown phase-rotation can be treated as an unknown quadrature-displacement, considerably simplifying the problem. The problem that we address in our paper is one of general phase estimation, where the phase cannot be approximated by a fixed quadrature (see Sec. S3). Moreover we consider a phase which varies continuously and stochastically in time.

To attack this problem we use an adaptive estimation technique, and sophisticated data processing. We estimate the phase at any time with a mean square error $(15 \pm 4)\%$ below the coherent-state limit (CSL) — the best precision that can be achieved with classical (i.e. coherent) states of light. This is possible because we use nonclassical phase-squeezed light, so called because the phase uncertainty is squeezed below the CSL (*7–9, 19, 20*). Because of the Heisenberg uncertainty principle, phase squeezing necessarily results in anti-squeezing in the amplitude. This anti-squeezing plays a key role in our experiment, causing there to be an optimal degree of squeezing, unlike in gravitational wave detection where more squeezing is

15

better (*7–9*). Our observation of this optimal squeezing is a new phenomenon in quantum phase estimation.

# S2  Details of the experiment

Figure S1 shows the experimental setup in more detail than shown in Fig. 1C in the main text. Specifically we show an optical mode cleaner, acousto-optic modulators (AOMs) and a radio frequency (RF) source, all of which were omitted from Fig. 1C in the main text in the interests of clarity.

The signal $\varphi(t)$, which is the waveform to be estimated, is imposed upon a weak phase-squeezed state using an electro-optic modulator. Adaptive homodyne detection of the phase-modulated beam is used to form the phase estimate.

In this section we describe the generation of the weak phase-squeezed state, and then describe its characterization using static homodyne detection. We then describe the generation of the signal and the phase estimation procedure using adaptive homodyne detection.

## S2.1 Generation and measurement of phase-squeezed states

We use an 860 nm continuous-wave (CW) Titanium Sapphire laser in the experiment. A fraction of the laser beam is used as the local oscillator (LO) for homodyne detection, with a spatial-mode cleaner used to improve mode-matching of the LO to the phase-squeezed state. The phase-squeezed state is generated by a sub-threshold optical parametric oscillator (OPO) using a periodically poled KTiOPO$_4$ crystal (*27*). The OPO is pumped by the 430 nm output of the second-harmonic-generator (SHG) cavity. The AOMs are used to generate a pair of optical sidebands at $\pm 5$ MHz (*21*), which are equivalent to a weak coherent state. The optical sidebands, which are within the 13 MHz half-width-half-maximum bandwidth of the OPO, are then injected into the OPO.
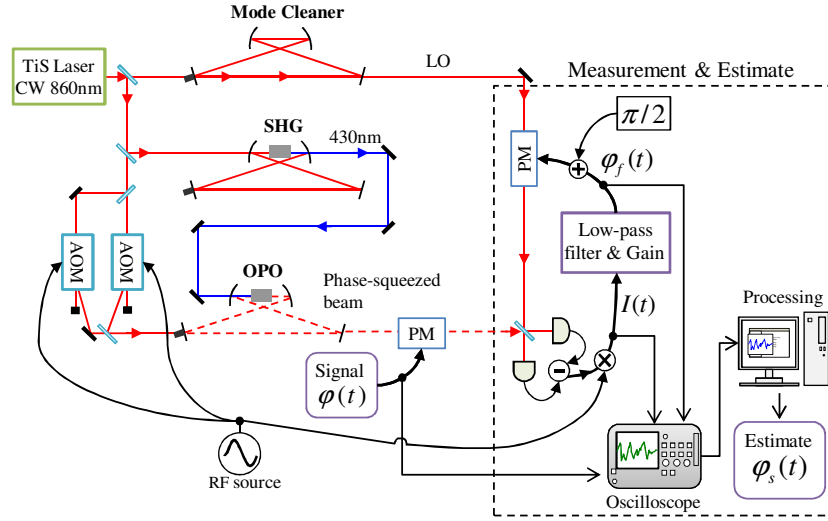
Figure S1: Experimental setup. The abbreviations are TiS: Titanium Sapphire, CW: continuous-wave, AOM: acousto-optic modulator, PM: phase modulator, SHG: second-harmonic-generation, OPO: optical parametric oscillator, RF source: radio frequency (5 MHz) source.

The relative phase between squeezing and the 5 MHz sidebands must be locked to ensure that a phase-squeezed beam is generated. For that purpose, we tap off 1% of the OPO output beam to be used as a signal to lock the relative phase between the squeezing and the sidebands. Since the OPO output power is extremely low (on the order of pW), we use another homodyne detector to measure the tapped beam (not shown in Fig. S1). The method is similar to that used in Ref. (*27*).

In addition to the deliberately-imposed phase modulation signal $\varphi(t)$, the experiment is subject to low-frequency phase disturbances from the environment. We use a low-frequency, low-gain feedback loop to suppress the environmental disturbances. When we use static homodyne detection to characterize the squeezing (anti-squeezing) level of the phase-squeezed states before performing the phase estimation experiment, we use only this classical feedback loop to lock the relative phase between the LO and the phase-squeezed beams.

With a LO beam power of 2 mW, the shot noise level of the homodyne detector is 10.4 dB

higher than the circuit noise level (that is, the ratio of the measured shot noise to circuit noise is $S = 11.0$ ). The homodyne visibility is $\xi = 0.988$, the optical transmission is $\rho = 0.97$, and the quantum efficiency of each of the photo-diodes is $\zeta = 0.99$. The overall efficiency of the homodyne detection is thus calculated as $\eta = \xi^2 \rho \zeta (S - 1)/S = 0.85$ (*28*).

Figure S2 shows static homodyne measurements of the phase-squeezed state when the power of the OPO pump beam is 180 mW. Fig. S2 shows the noise spectrum as measured with a radio frequency spectrum analyzer in the region around the 5 MHz sidebands. As we have mentioned previously, these are static homodyne measurements of the phase-squeezed state. So for Fig. S2 we do not apply phase modulation signal $\varphi(t)$ to the phase-squeezed beam, and do not use the phase-tracking loop. The residual peak at around 5 MHz in trace (ii) comes from fluctuations of the classical locking loop. Compared to our signal bandwidth ($\lambda/2\pi = 9.4$ kHz), the bandwidth of the squeezing is broad enough that we may assume that the squeezing is broadband over the frequency range of interest.

In our experiment, we use pump beam powers of 0, 30, 80 and 180 mW. Squeezing and anti-squeezing levels around 5 MHz are shown in Fig. S3. The solid line is a theoretical curve fitted to the experimental results with the overall loss $l_{\mathrm{sq}} = 1 - \eta(1 - l_{\mathrm{s}})$ (comprising the homodyne detection efficiency, $\eta$, and the loss from the OPO and phase modulator, $l_{\mathrm{s}}$) being a free parameter. We define the measured squeezing and anti-squeezing levels as $R_- := e^{-2r_m}$ and $R_+ := e^{2r_p}$. With an overall loss $l_{\mathrm{sq}}$ and a pure squeezing parameter $r$, the measured squeezing and anti-squeezing levels can be written as (*29*),

$$
\begin{aligned}
R_- &:= e^{-2r_m} = (1 - l_{\mathrm{sq}})e^{-2r} + l_{\mathrm{sq}}, \\
R_+ &:= e^{2r_p} = (1 - l_{\mathrm{sq}})e^{2r} + l_{\mathrm{sq}}.
\end{aligned}
\tag{4}
$$

Hence we obtain the following equation,

$$
R_+ = \frac{(1 - l_{\mathrm{sq}})^2}{R_- - l_{\mathrm{sq}}} + l_{\mathrm{sq}}.
\tag{5}
$$

In Fig. S3, we obtain the overall loss $l_{\text{sq}} = 0.33$ which mainly comes from the OPO and phase modulator.
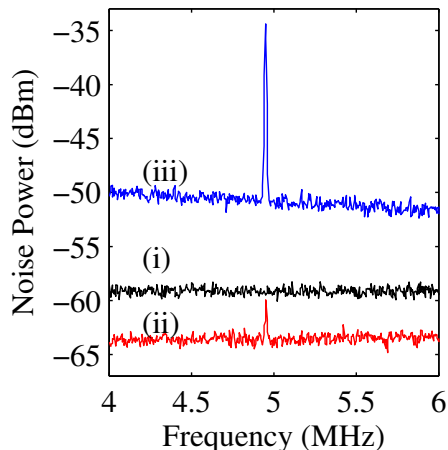


Figure S2: Characterizing the phase squeezing. The data is recorded with a spectrum analyzer. Resolution and video bandwidths are 10 kHz and 300 Hz, respectively. Trace (i) is shot noise level. Traces (ii) and (iii) are the squeezing and anti-squeezing levels where the LO phase is locked to those quadratures.
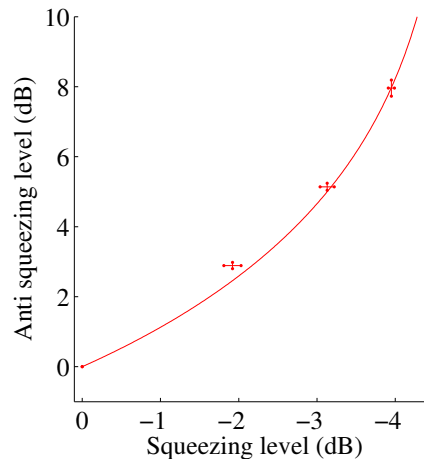
Figure S3: Anti-squeezing level versus squeezing level. Solid curve is fitted to the experimental data with loss being a free parameter. The overall loss $l_{\text{sq}}$ is obtained as 0.33.

## S2.2 Phase Tracking

The signal waveform to be estimated is generated by a random signal source and a low-pass filter with a cutoff frequency of $\lambda/2\pi = 9.4$ kHz. The signal source is a digital signal generator (AFG3021, Tektronix), which generates white Gaussian noise (3dB bandwidth is 25 MHz). The signal $\varphi(t)$ is imposed upon the phase-squeezed beam using an electro-optic modulator (LM0202, Linos). We perform homodyne detection on this phase-modulated beam. After homodyne detection of the phase-modulated beam, the homodyne current is demodulated at 5 MHz and then goes to the tracking filter (another low-pass filter with a cutoff frequency $\lambda/2\pi$ according to the Kalman filter theory). The output of the tracking filter gives the filtered

estimate $\varphi_f(t)$, which is then shifted by $\pi/2$ and applied to the phase of the LO beam with another electro-optic modulator (waveguide modulator, EOSPACE ). The gain of the phase-lock loop is optimized to give the maximum sensitivity. As explained previously, an additional low-frequency, low gain feedback loop is used to isolate the experiment from environmental disturbances.

We record $\varphi(t)$, $\varphi_f(t)$ and the demodulated homodyne current $I(t)$ using an oscilloscope with a sampling rate of 100 MHz. We calculate the mean square error from 2 ms of data, comprising $2 \times 10^5$ data points. To calculate the average and the standard deviation of the mean square error, we repeat the measurement 15 times. Note that each the correlation time of the noise signal $\varphi(t)$ is approximately $\lambda^{-1}$, and so $2 \times 10^5$ data points have effectively $2\,\mathrm{ms}/\lambda^{-1} \sim 10^2$ independent data points.

## S3  Phase estimation with broadband squeezing

In this section we describe phase estimation with broadband squeezing. As in the main text, the phase variation $\varphi(t)$ obeys the following,

$$\varphi(t) = \sqrt{\kappa} \int_{-\infty}^{t} e^{-\lambda(t-s)} dV(s), \tag{6}$$

where $dV(s)$ is a Wiener increment which satisfies $\langle dV(s)dV(s') \rangle = \delta(s - s')(dt)^2$, $\lambda$ is the bandwidth of the phase noise $\varphi$, and $\kappa/2\lambda$ is the mean square variation of $\varphi$. Note that in our experiment this variation is much larger than mean square error with which we can estimate $\varphi$. The phase-modulated beam is measured by homodyne detection. The LO phase $\Phi(t)$ is adaptively controlled to be $\Phi(t) := \varphi_f(t) + \pi/2$, where $\varphi_f(t) \approx \varphi(t)$ is our filtered estimate. In our experiment the squeezing bandwidth is many MHz, which is much larger than the bandwidth $\lambda$ of the phase noise. Hence we can approximate the squeezing noise as white. Using the

normalization of (*19*), the homodyne output current is

$$I(t)dt \;=\; 2\,|\alpha|\sin\left[\varphi(t) - \varphi_f(t)\right]dt + \sqrt{R_{\text{sq}}(t)}dW(t), \tag{7}$$

$$R_{\text{sq}}(t) \;=\; \sin^2\left[\varphi(t) - \varphi_f(t)\right]e^{2r_p} + \cos^2\left[\varphi(t) - \varphi_f(t)\right]e^{-2r_m}, \tag{8}$$

where $|\alpha|^2$ is the photon flux of the coherent component of the beam (i.e. photons per unit time), $dW(t)$ is a Wiener increment representing quantum white noise, satisfying $\langle dW(s)dW(s')\rangle = \delta(s - s')(dt)^2$, and the amplitude of this noise $R_{\text{sq}}$ is determined by the squeezing ($r_m \geq 0$) and anti-squeezing ($r_p \geq 0$) parameters defined earlier. The squeezing factor $R_{\text{sq}}(t)$ is time-dependent because of tracking error $\varphi(t) - \varphi_f(t)$, the unknown and varying misalignment between the squeezing ellipse and the local oscillator phase.

We denote the tracking error by $\Delta_f(t) = \varphi(t) - \varphi_f(t)$, and its stationary ensemble average by $\sigma_f^2 = \langle\Delta_f^2\rangle_{\text{ss}}$. Because $\sigma_f \ll 1$ in our experiment we can approximate $I(t)dt$ by a 2nd order expansion in $\Delta_f(t)$. This gives

$$I(t)dt \;\simeq\; 2\,|\alpha|\,\Delta_f(t)dt + \sqrt{R_{\text{sq}}}dW(t), \tag{9}$$

$$R_{\text{sq}}(t) \;\simeq\; \Delta_f^2(t)e^{2r_p} + [1 - \Delta_f^2(t)]e^{-2r_m}. \tag{10}$$

The approximations given here help to illustrate the difference between phase estimation and phase sensing. In phase sensing, the variation in the system phase $\varphi(t)$ is sufficiently small that $I(t)dt$ can be linearised in $\varphi(t)$. In phase estimation, this is not the case and a better approximation is needed. Here the filtered estimate $\varphi_f(t)$ is sufficiently good that the sine function can be linearised in $\Delta_f(t)$. (This is why feedback is vital in phase estimation.) Furthermore, in the case where there is squeezing, there is a difference in the treatment of the squeezing term $R_{\text{sq}}(t)$. In phase sensing, the variation in the system phase is sufficiently small that one may simply take $R_{\text{sq}}(t) \simeq e^{-2r_m}$. One can simply treat the phase-rotation as a displacement of the squeezed quadrature, and not consider the anti-squeezing. In contrast,

21

for phase estimation (even with feedback) the variation of the system phase is sufficiently large that one needs a better approximation for $R_{sq}(t)$. Because the filtered estimate is not perfectly accurate, there is a significant contribution to the error from the anti-squeezed quadrature.

A further approximation is to replace $\Delta_f^2(t)$ by its stationary average $\sigma_f^2$, which is independent of time. (This approximation is justified below by comparing the time scale of its variation to that of the subsequent filter.) The normalized homodyne output current can then be written as

$$I(t)dt \simeq 2|\alpha|\Delta_f(t)dt + \sqrt{\bar{R}_{sq}}dW(t), \tag{11}$$

$$\bar{R}_{sq} = \sigma_f^2 e^{2r_p} + (1 - \sigma_f^2)e^{-2r_m}, \tag{12}$$

where $\bar{R}_{sq}$ is an effective (time-independent) squeezing factor which takes into account the tracking error in an average sense. A measurement better than the coherent-state limit is expected to be possible when this $\bar{R}_{sq}$ is less than $1$.

For an equation of the form of Eq. 11, it is known that the optimal causal estimate (as required for tracking) is provided by the Kalman filter (22):

$$\varphi_f(t) = \Gamma \int_{-\infty}^{t} e^{-\lambda(t-s)} \frac{I(s)}{2|\alpha|} ds, \tag{13}$$

where $\Gamma$ is the Kalman gain. Before going further, we return to the approximation in Eq. 12, where we replace the time-varying $\Delta_f^2(t)$, with its average $\sigma_f^2$. The characteristic time of the Kalman filter is $\lambda^{-1}$; that is, the homodyne output current $I(t)$ is, roughly speaking, averaged over the period of $\lambda^{-1}$. Thus if the variation of $\Delta_f^2(t)$ is rapid compared to $\lambda^{-1}$, it is justified to replace the time-varying $\Delta_f^2(t)$ by the average $\sigma_f^2$. Figure S4 shows a typical experimental trace of the temporal variation of the squared filtered estimate error $\Delta_f^2(t) = [\varphi(t) - \varphi_f(t)]^2$, and also its auto-correlation function. The experimental conditions are the same as Fig. 2 in the main text. The variation of the squared filtered estimate error has a correlation time around $2\mu$s.

The Kalman filter characteristic time is calculated as $\lambda^{-1} \simeq 17\mu s$ which is much larger than the correlation time of the squared filtered estimate error. Therefore our approximation in Eq. 12 turns out to be appropriate.
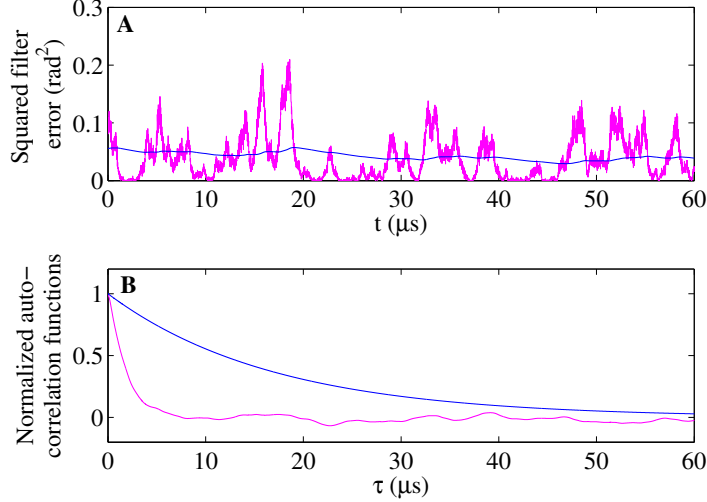


Figure S4: (**A**) Squared error of the filtered estimate $\Delta_f^2$ under the same conditions as Fig. 2 in the main text. The purple trace is $\Delta_f^2$, and the blue trace is that after the Kalman filter. (**B**) Normalized auto-correlation $\langle [\Delta_f(t)]^2, [\Delta_f(t+\tau)]^2 \rangle$ as measured in the experiment (purple curve), and theoretical curve for the Kalman filter of white noise (blue curve).

This separation of time scales is a consequence of the fact that our experiment was performed in the limit $\Gamma^{-1} \ll \lambda^{-1}$ (for the traces in Fig. S4, $\Gamma^{-1} \approx 4\mu s$). This can be seen by considering the differential equation for $\Delta_f(t)$. First, combining Eqs. 11 and 13 gives

$$\varphi_f(t) = \Gamma \int_{-\infty}^{t} e^{-\lambda(t-s)} \left[ \Delta_f(s)ds + \frac{\sqrt{R_{\text{sq}}}}{2\,|\alpha|}dW(s) \right]. \tag{14}$$

Combining this with Eq. 6 gives

$$d\Delta_f \;=\; d\varphi(t) - d\varphi_f(t) \tag{15}$$

$$=\; -(\lambda+\Gamma)\Delta_f(t)dt + \sqrt{\kappa}dV(t) - \frac{\Gamma\sqrt{R_{\text{sq}}}}{2\,|\alpha|}dW(t). \tag{16}$$

23

Thus, under the above approximation, $\Delta_f$ has a correlation time of $(\lambda+\Gamma)^{-1}$, and $\Delta_f^2$ a correlation time of half that. For $\Gamma \gg \lambda$, the correlation time for $\Delta_f^2$ is much less than that of the filter, and hence on the time scale of the filter $\Delta_f^2$ may be replaced by its ensemble average. More fundamentally, this difference in time scales comes because we have conducted our experiment in the regime where $\epsilon \equiv \sqrt{\lambda^2 \bar{R}_{\rm sq}/(4\,|\alpha|^2\,\kappa)}$ is small, as in this regime $\Gamma \simeq \lambda/\epsilon$ (see below).

Even if we do not replace $R_{\rm sq}$ (10) by $\bar{R}_{\rm sq}$ (12), we find the same result for $\sigma_f^2$. This follows from the differential equation for $\left\langle \Delta_f^2 \right\rangle$ using Eq. 10, namely

$$d\left\langle \Delta_f^2 \right\rangle = -2(\lambda+\Gamma)\left\langle \Delta_f^2 \right\rangle dt + \kappa dt - \frac{\Gamma^2 \left\langle \Delta_f^2 e^{2r_p} + (1-\Delta_f^2)e^{-2r_m} \right\rangle}{4\,|\alpha|^2}dt. \tag{17}$$

This gives a linear equation for $\sigma_f^2 = \langle \Delta_f^2 \rangle_{\rm ss}$ which is identical to the equation which would be obtained from Eq. 16 in which $\bar{R}_{\rm sq}$ contains $\sigma_f^2$ already. This means that our theory models the experiment for any value of $\epsilon$ (provided $\sigma_f^2$ remains small), but our argument for the optimality of the Kalman filter, and the corresponding smoothed estimator, holds only for $\epsilon$ small. This is reasonable for our experiment, in which $\epsilon \leq 0.2$.

Returning to Eq. 14, it is obvious that phase estimation using the phase-squeezed beam with the amplitude $|\alpha|$ and squeezing factor $\bar{R}_{\rm sq}$ is mathematically equivalent to that using a coherent beam with the amplitude $|\alpha|/\sqrt{\bar{R}_{\rm sq}}$. (Note that this equivalence holds in terms of the estimate precision, not in terms of photon number.) Therefore we can obtain the Kalman gain $\Gamma$ and the mean square error $\sigma_f^2$ by formally replacing $|\alpha|$ in Ref. (22) with $|\alpha|/\sqrt{\bar{R}_{\rm sq}}$, giving

$$\Gamma = -\lambda + \sqrt{\lambda^2 + 4\kappa\,|\alpha|^2/\bar{R}_{\rm sq}}, \tag{18}$$

$$\sigma_f^2 = \frac{\lambda\bar{R}_{\rm sq}}{4\,|\alpha|^2}\left(\sqrt{1 + \frac{4\kappa\,|\alpha|^2}{\lambda^2\bar{R}_{\rm sq}}} - 1\right). \tag{19}$$

In the limit $\epsilon \ll 1$, the optimal gain is $\Gamma \approx \lambda/\epsilon$, and $\sigma_f^2 \approx \sqrt{\kappa\bar{R}_{\rm sq}}/(2\,|\alpha|)$. All of these equations are still implicit because $\bar{R}_{\rm sq}$ (12) is a function of $\sigma_f^2$. Solving the implicit equation

for $\sigma_f^2$ gives explicitly, and exactly (i.e. without assuming $\epsilon \ll 1$)

$$\Gamma = -\lambda - \frac{\kappa\left(e^{2r_p} - e^{-2r_m}\right)}{2e^{-2r_m}} + \frac{1}{2}\sqrt{\left(\frac{\kappa\left(e^{2r_p} - e^{-2r_m}\right)}{e^{-2r_m}} + 2\lambda\right)^2 + \frac{16|\alpha|^2\kappa}{e^{-2r_m}}}, \quad (20)$$

$$\sigma_f^2 = \frac{1}{8|\alpha|^2 + 4\lambda(e^{2r_p} - e^{-2r_m})}\Bigg[-2\lambda e^{-2r_m} + \kappa(e^{2r_p} - e^{-2r_m})$$

$$+\sqrt{[2\lambda e^{-2r_m} + \kappa(e^{2r_p} - e^{-2r_m})]^2 + 16|\alpha|^2\kappa e^{-2r_m}}\,\Bigg]. \quad (21)$$

As in Ref. (*21*), we can apply the smoothing technique of Ref. (*22*) for further improvement. The smoothed estimate $\varphi_s(t)$ and mean square error $\sigma_s^2 = \left\langle [\varphi(t) - \varphi_s(t)]^2 \right\rangle$ are

$$\varphi_s(t) = (2\lambda + \Gamma)\int_t^\infty e^{-(\lambda+\Gamma)(s-t)}\varphi_f(s)ds, \quad (22)$$

$$\sigma_s^2 = \frac{\kappa}{2\sqrt{4\kappa\,|\alpha|^2/\bar{R}_{\text{sq}} + \lambda^2}}. \quad (23)$$

These formulae are obtained by formally replacing $|\alpha|$ in Ref. (*22*) with $|\alpha|/\sqrt{\bar{R}_{\text{sq}}}$. In the limit $\epsilon \ll 1$, we obtain the simple relation $\sigma_s^2 \approx \sigma_f^2/2$ as in Ref. (*21*).

## S4 Optimum squeezing level for phase tracking

In this section, we discuss the optimum squeezing level for the phase tracking, as quantified by the mean square error in the smoothed estimate.

From Eq. 12 and 23, it is obvious that highly squeezed states would not improve the phase tracking if there were too much anti-squeezing noise. Quantum mechanics does not allow us to squeeze a physical quantity without anti-squeezing the canonically conjugate counterpart. Hence there must be an optimum squeezing level for the phase estimation problem. Figure S5 shows the smoothed mean square error $\sigma_s^2$ as a function of both squeezing and anti-squeezing levels, with other parameters ($\alpha$, $\kappa$, $\lambda$) fixed. In this figure the lower-right half is forbidden due to the uncertainty principle. Even for a pure squeezed state (at the boundary of the forbidden region), the minimum mean square error is obtained for an optimum squeezing level of about

25

7 dB. In practice, with finite loss, the squeezed state is not pure, and it is not possible to attain the boundary with the forbidden region. The minimum anti-squeezing for a given squeezing will follow a curve away from the boundary, according to Eq. 5. The curve for the loss in our experiment is shown in Figure S5. The qualitative behavior is the same as in the pure state case. There is an optimal squeezing level, and increasing the squeezing beyond this will increase the mean square error.
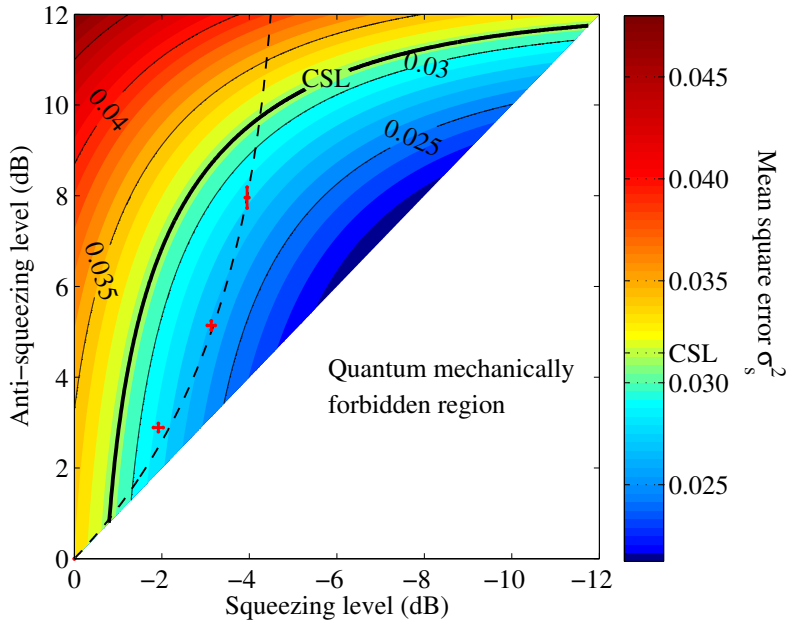


Figure S5: Mean square error of the smoothed estimate, as a function of both squeezing and anti-squeezing levels, calculated from Eq. 23. Other parameters ($\alpha$, $\kappa$, $\lambda$) are fixed and the same as Fig. 3 in the main text. The black dashed line and red crosses are the predicted and experimentally observed squeezing and anti-squeezing levels in our setup (see Fig. S3). CSL stands for the coherent-state limit.

The optimum squeezing level varies with parameters $\alpha$, $\kappa$ and $\lambda$. From Eq. 12, it would be obvious that the optimum squeezing level is higher if the mean square error of the filtered estimate $\sigma_f^2$ is smaller. This is because the anti-squeezing noise has less effect if $\sigma_f^2$ is small. The mean square error $\sigma_f^2$ would be reduced by increasing $\alpha$, or $\lambda$, or by reducing $\kappa$. Note that

the mean square variation of signal $\varphi$ is $\kappa/2\lambda$, and the estimate error $\sigma_f^2$ is smaller with smaller signal variation (smaller $\kappa$ or larger $\lambda$). Therefore, the optimum squeezing level is higher if $\alpha$ or $\lambda$ is larger or $\kappa$ is smaller (see Fig. S6).
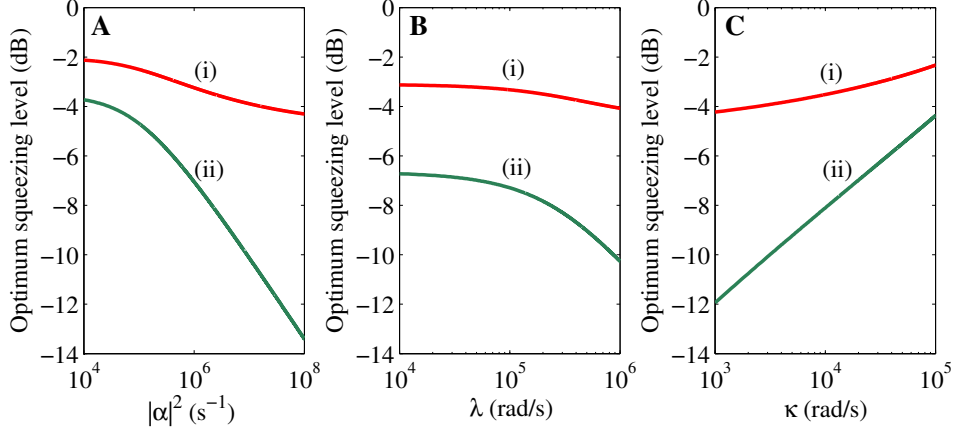


Figure S6: Optimum squeezing levels as a function of (**A**) $|\alpha|^2$, (**B**) $\lambda$ and (**C**) $\kappa$. The optimum squeezing levels are calculated by minimizing $\sigma_s^2$ of Eq. 23. The parameters other than the varying parameter are fixed and the same as Fig. 3 in the main text (e.g., in (**A**), $\kappa$ and $\lambda$ are fixed and the same as Fig. 3 in the main text.). Traces (i) are for squeezed beams in our experimental setup (i.e. including loss). Traces (ii) are for pure squeezed beams.

# S5  Squeezing bandwidth and photon flux

In this section, we derive the photon flux due to squeezing, as this is required (see Sec. S7) for a complete accounting of the resources used in phase estimation. As we shall show shortly, infinite-bandwidth squeezing leads to an infinite photon flux. Therefore, finite-bandwidth squeezing must be considered explicitly.

Let us consider a squeezing spectrum $R_-(\Omega)$ and anti-squeezing spectrum $R_+(\Omega)$ of the

standard form (*27*)

$$R_{\pm}(\Omega) \;=\; 1 + (R_{\pm} - 1)\frac{[(1 \mp x)\Delta\Omega_0]^2}{\Omega^2 + [(1 \mp x)\Delta\Omega_0]^2}, \tag{24}$$

$$x \;=\; \frac{R_+ - R_- - 2\sqrt{(1 - R_-)(R_+ - 1)}}{R_+ + R_- - 2}, \tag{25}$$

where $R_{\pm}$ are squeezing and anti-squeezing levels at the center frequency ($\Omega = 0$). In the case of the OPO, $2\Delta\Omega_0$ and $x$ correspond to the cavity's decay rate and the normalized pump amplitude respectively. The difference between the bandwidths for squeezing $(1 + x)\Delta\Omega_0$ and anti-squeezing $(1 - x)\Delta\Omega_0$ ensures that $R_+(\Omega)R_-(\Omega) = 1$ for all $\Omega$ when the squeezed state is pure.

The number of photons due to squeezing in a frequency mode $\Omega$ is given as (*29*),

$$n(\Omega) = \frac{1}{4}\left[R_+(\Omega) + R_-(\Omega)\right] - \frac{1}{2}. \tag{26}$$

The photon flux per unit time $\mathcal{N}_{\mathrm{sq}}$ due to the squeezing is given by integrating $n(\Omega)$ over $\Omega$ (*30*),

$$\mathcal{N}_{\mathrm{sq}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} n(\Omega)d\Omega = \frac{1}{4}\left[(R_+ - 1)(1 - x) + (R_- - 1)(1 + x)\right]\frac{\Delta\Omega_0}{2}. \tag{27}$$

Note that, as mentioned previously, the photon flux diverges for infinite-bandwidth squeezing.

## S6  Phase estimation with  finite-bandwidth squeezing

In this section, we derive mean square error of the phase estimate when the squeezing bandwidth is finite. This will be used in Sec. S7 to calculate the (small) effect that restricting the bandwidth of our squeezing would have had.

For finite-bandwidth squeezing the equations 11 and 12 for the homodyne output current must be modified to

$$I(t)dt \;\simeq\; 2|\alpha|\left[\varphi(t) - \tilde{\varphi}_f(t)\right]dt + dY(t), \tag{28}$$

$$\langle dY(s)dY(s')\rangle \;=\; dsds'\left[\tilde{\sigma}_f^2 X_+(s - s') + (1 - \tilde{\sigma}_f^2)X_-(s - s')\right], \tag{29}$$

where $X_\pm(s - s')$ are auto-correlation functions for the squeezed $(-)$ and anti-squeezed $(+)$ quadratures, and the tilde is used to distinguish particular finite-bandwidth properties from the previously calculated infinite-bandwidth quantities (with no tilde). $X_\pm(s - s')$ are calculated by the inverse Fourier transform of the squeezing and anti-squeezing spectrum $R_\pm(\Omega)$,

$$
\begin{aligned}
X_\pm(s - s') &= \frac{1}{2\pi} \int_{-\infty}^{\infty} R_\pm(\Omega) e^{i\Omega(s-s')} d\Omega \\
&= \delta(s - s') + (R_\pm - 1) \frac{(1 \mp x)\Delta\Omega_0}{2} e^{-(1\mp x)\Delta\Omega_0 |s-s'|}.
\end{aligned} \tag{30}
$$

From Eqs. 13 and 28, and assuming a filter of the same form as above, the estimator used for tracking is

$$
\tilde{\varphi}_f(t) = \tilde{\Gamma} \int_{-\infty}^{t} e^{-\lambda(t-s)} \frac{I(s)}{2|\alpha|} ds = \tilde{\Gamma} \int_{-\infty}^{t} e^{-\lambda(t-s)} \left\{ [\varphi(s) - \tilde{\varphi}_f(s)] \, ds + \frac{dY(s)}{2\,|\alpha|} \right\}, \tag{31}
$$

where the optimal Kalman gain $\tilde{\Gamma}$ can be found numerically. In order to solve this equation, we transform it into differential form as

$$
\begin{aligned}
d\tilde{\varphi}_f(t) &= -\lambda\tilde{\varphi}_f(t)dt + \tilde{\Gamma} \left\{ [\varphi(t) - \tilde{\varphi}_f(t)] \, dt + \frac{dY(t)}{2\,|\alpha|} \right\} \\
&= -(\lambda + \tilde{\Gamma})\tilde{\varphi}_f(t)dt + \tilde{\Gamma} \left[ \varphi(t)dt + \frac{dY(t)}{2\,|\alpha|} \right].
\end{aligned} \tag{32}
$$

Then we transform it into integral form again,

$$
\tilde{\varphi}_f(t) = \tilde{\Gamma} \int_{-\infty}^{t} e^{-(\lambda+\tilde{\Gamma})(t-s)} \left[ \varphi(s)ds + \frac{dY(s)}{2\,|\alpha|} \right]. \tag{33}
$$

From this, we can see that the noise term $dY(t)$ is low-pass filtered with a cutoff (angular) frequency $(\lambda + \tilde{\Gamma})$. Thus we could consider that the bandwidth of the estimator is $(\lambda + \tilde{\Gamma})$.

We calculate the mean square error $\tilde{\sigma}_f^2$ as,

$$
\begin{aligned}
\tilde{\sigma}_f^2 &= \left\langle [\varphi(t) - \tilde{\varphi}_f(t)]^2 \right\rangle \\
&= \left\langle \left\{ \tilde{\Gamma} \int_{-\infty}^{t} e^{-(\lambda+\tilde{\Gamma})(t-s)} \left[ \frac{\lambda+\tilde{\Gamma}}{\tilde{\Gamma}} \varphi(t) - \varphi(s) \right] ds - \tilde{\Gamma} \int_{-\infty}^{t} e^{-(\lambda+\tilde{\Gamma})(t-s)} \frac{dY(s)}{2|\alpha|} \right\}^2 \right\rangle \\
&= \tilde{\Gamma}^2 \int_{-\infty}^{t} ds_1 \int_{-\infty}^{t} ds_2 e^{-(\lambda+\tilde{\Gamma})(2t-s_1-s_2)} \left\langle \left[ \frac{\lambda+\tilde{\Gamma}}{\tilde{\Gamma}} \varphi(t) - \varphi(s_1) \right] \left[ \frac{\lambda+\tilde{\Gamma}}{\tilde{\Gamma}} \varphi(t) - \varphi(s_2) \right] \right\rangle \\
&\quad + \frac{\tilde{\Gamma}^2}{4|\alpha|^2} \int_{-\infty}^{t} \int_{-\infty}^{t} e^{-(\lambda+\tilde{\Gamma})(2t-s_1-s_2)} \left\langle dY(s_1) dY(s_2) \right\rangle .
\end{aligned}
\tag{34}
$$

Here $\langle \varphi(s)\varphi(s') \rangle$ is calculated from Eq. 6 as,

$$
\langle \varphi(s)\varphi(s') \rangle = \frac{\kappa}{2\lambda} e^{-\lambda|s-s'|}.
\tag{35}
$$

Now $\tilde{\sigma}_f^2$ can be given from Eqs. 29, 30, 34 and 35,

$$
\begin{aligned}
\tilde{\sigma}_f^2 &= \frac{\kappa}{2(\lambda+\tilde{\Gamma})} + \frac{1}{8|\alpha|^2} \frac{\tilde{\Gamma}^2}{\tilde{\Gamma}+\lambda} \\
&\quad \times \left\{ 1 + \tilde{\sigma}_f^2(R_+ - 1)\frac{h_+}{h_+ + 1} + (1-\tilde{\sigma}_f^2)(R_- - 1)\frac{h_-}{h_- + 1} \right\},
\end{aligned}
\tag{36}
$$

$$
h_\pm = \frac{(1 \mp x)\Delta\Omega_0}{\lambda + \tilde{\Gamma}}.
\tag{37}
$$

The effect of finite-bandwidth squeezing is characterized by $h_\pm$, which is the ratio of the squeezing (anti-squeezing) bandwidth $(1 \mp x)\Delta\Omega_0$ to the estimator bandwidth $(\lambda + \tilde{\Gamma})$. In the limit $\Delta\Omega_0 \to \infty$ we find from Eq. 36 that $\tilde{\sigma}_f^2$ obtains its minimum value of $\sigma_f^2$ when $\tilde{\Gamma} = \Gamma$, as expected.

The optimal smoothed estimate with finite-bandwidth squeezing $\tilde{\varphi}_s(t)$ is similarly found to be

$$
\tilde{\varphi}_s(t) = \left(2\lambda + \tilde{\Gamma}\right) \int_{t}^{\infty} e^{-(\lambda+\tilde{\Gamma})(s-t)} \tilde{\varphi}_f(s) ds,
\tag{38}
$$

and the mean square error of the smoothed estimate is

$$
\begin{aligned}
\tilde{\sigma}_s^2 &= \frac{\kappa(2\lambda^2 + \tilde{\Gamma}^2 + 2\lambda\tilde{\Gamma})}{4(\lambda + \tilde{\Gamma})^3} + \frac{\tilde{\Gamma}^2(2\lambda + \tilde{\Gamma})^2}{16|\alpha|^2(\lambda + \tilde{\Gamma})^3} \\
&\quad \times \left\{ 1 + \tilde{\sigma}_f^2(R_+ - 1)\frac{h_+(h_+ + 2)}{(h_+ + 1)^2} + (1 - \tilde{\sigma}_f^2)(R_- - 1)\frac{h_-(h_- + 2)}{(h_- + 1)^2} \right\}.
\end{aligned}
\tag{39}
$$

Similarly to that for the filtered estimate, the effect of finite-bandwidth squeezing is characterized by $h_\pm$. Again, in the limit $\Delta\Omega_0 \to \infty$ we find from Eq. 39 that $\tilde{\sigma}_s^2$ obtains its minimum value of $\sigma_s^2$ when $\tilde{\Gamma} = \Gamma$.

## S7 Effective squeezing bandwidth and photon flux

In our experiment, we use narrow optical sidebands as our signal. The frequency range of squeezing which is exploited, in the vicinity of the sideband frequency, is thus much narrower than the OPO's bandwidth. In other words, most of the photon flux produced by the squeezing plays no part in the current experiment. Counting the entire photon flux of the beam would grossly misrepresent the photon flux resource used by our phase tracking algorithm. To reasonably represent the photons in the squeezing (in addition to those in the coherent field) we must consider the photons in a relatively narrow band in the vicinity of the sideband frequency. That is, we must define the *effective* squeezing (and anti-squeezing) bandwidths.

Once again we model the spectrum of the squeezing and anti-squeezing by Eq. 24, with $\Omega = 0$ denoting the sideband frequency. As before, $R_\pm$ denotes the squeezing (anti-squeezing) levels at the sideband frequency, but now we replace $\Delta\Omega_0$ by an effective bandwidth $\Delta\Omega_{\text{eff}}$ to model narrow-band squeezing. The effective bandwidth is the narrowest bandwidth of squeezing that could have been used without substantially affecting the experimental results obtained. A reasonable guess would be to set the effective bandwidth somewhat wider than the estimator bandwidth. This is because the squeezing within the estimator bandwidth can effectively enhance the estimate, whereas the squeezing outside of the estimator bandwidth is mostly wasted.

Indeed, it turns out that defining $\Delta\Omega_{\text{eff}} = 2(\lambda + \Gamma)$ is appropriate in the sense that $\tilde{\sigma}_s^2$ is little different from $\sigma_s^2$ in the regime of our experiment.

From now on, we restrict to the experimental parameters as shown in Fig. 4 of the main text. The squared coherent amplitude $|\alpha|^2$ ranges from $10^6$ to $10^7 s^{-1}$, and the other parameters are fixed as $\kappa = 1.9 \times 10^4$ rad/s, $\lambda = 5.9 \times 10^4$ rad/s, $R_- = 0.479$, and $R_+ = 3.09$ (see the main text). The normalized pump amplitude $x$ is calculated as $x = 0.33$ from Eq. 25. In Fig. S7, we plot the mean square error of the phase estimate as a function of $|\alpha|^2$ (remember that $\Gamma$ is a function of $|\alpha|^2$ in Eq. 20, and hence $\Delta\Omega_{\text{eff}}$ is too). The green line indicates the mean square error $\sigma_s^2$ for broadband squeezing ($\Delta\Omega_{\text{eff}} \to \infty$). The red line is the mean square error $\tilde{\sigma}_s^2$ with $\Delta\Omega_{\text{eff}} = 2(\lambda + \Gamma)$. The difference between $\sigma_s^2$ and $\tilde{\sigma}_s^2$ is less than 3%, which is smaller than the experimental error bars which are $\sim 10\%$.

By setting $\Delta\Omega_{\text{eff}} = 2(\lambda + \Gamma)$, we can now calculate photon flux of the squeezing $\mathcal{N}_{\text{eff}}^{\text{sq}}$ from Eq. 27, and obtain the total effective photon flux $\mathcal{N}_{\text{eff}} = |\alpha|^2 + \mathcal{N}_{\text{eff}}^{\text{sq}}$. The contribution of $\mathcal{N}_{\text{eff}}^{\text{sq}}$ to the total photon flux is at most 7% under our experimental conditions. Figure S8 shows the mean square error versus the effective photon flux $\mathcal{N}_{\text{eff}}$ under the same conditions as Fig. 4 in the main text. We also replot the experimental data from that figure. To make an absolutely fair comparison, we correct the experimental mean square errors for broadband squeezing to the expected (i.e. slightly larger) values for the effective bandwidth. The key point is that even taking into account the photon flux due to squeezing in the effective bandwidth around the sidebands, the accuracy of our phase estimates surpasses the ideal (lossless) coherent-state limit as a function of total photon flux over the whole range of the experiment.
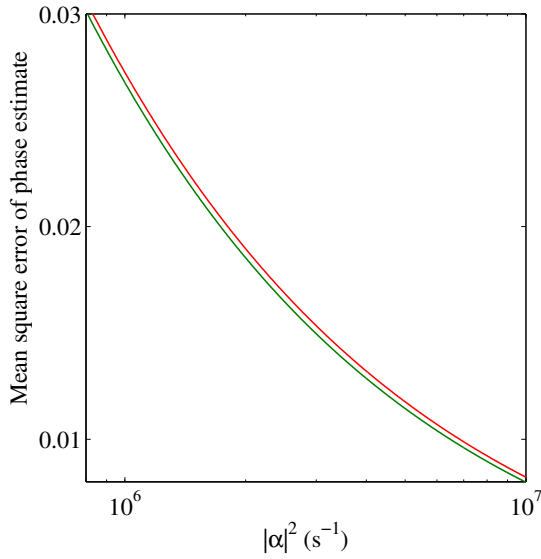
Figure S7: Theoretical curves for the mean square error in the smoothed estimate versus the photon flux $|\alpha|^2$ due to the coherent amplitude. Experimental conditions ($\kappa$, $\lambda$, and $R_{\pm}$) are same as Fig. 4 in the main text. The green line is $\sigma_s^2$ for broadband squeezing (as in the experiment). The red line is $\tilde{\sigma}_s^2$ for finite-bandwidth squeezing with $\Delta\Omega_{\text{eff}} = 2(\lambda + \Gamma)$, where $\Gamma$ is the feedback gain.
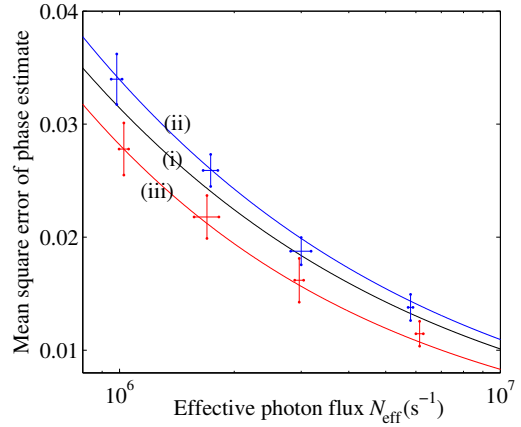
Figure S8: Mean square error of the smoothed estimate versus total effective photon flux $\mathcal{N}_{\text{eff}}$. Parameters are as in Fig. S7. Blue and red crosses are experimental data for coherent and squeezed beams respectively. The latter are corrected to the expected values for squeezing with bandwidth $\Delta\Omega_{\text{eff}}$. Trace (i) is the theoretical coherent-state limit. Trace (ii) is the predicted (i.e. with experimental imperfections) curve for coherent beams. Trace (iii) is the predicted curve for finite-bandwidth squeezed beams including all experimental imperfections.