

Quantum-enhanced stochastic phase estimation with the SU(1,1) interferometer

KAIMIN ZHENG,^{1,†} MINGHAO MI,^{1,†} BEN WANG,¹ LIANG XU,¹ LIYUN HU,² SHENGSHUAI LIU,³ YANBO LOU,³ JIETAI JING,^{3,4,5} AND LIJIAN ZHANG^{1,6}

¹National Laboratory of Solid State Microstructures, Key Laboratory of Intelligent Optical Sensing and Manipulation, College of Engineering and Applied Sciences, and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China

²Center for Quantum Science and Technology, Jiangxi Normal University, Nanchang 330022, China

³State Key Laboratory of Precision Spectroscopy, Joint Institute of Advanced Science and Technology, School of Physics and Electronic Science, East China Normal University, Shanghai 200062, China

⁴Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan 030006, China

⁵e-mail: jtjing@phy.ecnu.edu.cn

⁶e-mail: lijian.zhang@nju.edu.cn

Received 21 April 2020; revised 16 July 2020; accepted 2 August 2020; posted 4 August 2020 (Doc. ID 395682); published 29 September 2020

Quantum stochastic phase estimation has many applications in the precise measurement of various physical parameters. Similar to the estimation of a constant phase, there is a standard quantum limit for stochastic phase estimation, which can be obtained with the Mach–Zehnder interferometer and coherent input state. Recently, it has been shown that the stochastic standard quantum limit can be surpassed with nonclassical resources such as squeezed light. However, practical methods to achieve quantum enhancement in the stochastic phase estimation remain largely unexplored. Here we propose a method utilizing the SU(1,1) interferometer and coherent input states to estimate a stochastic optical phase. As an example, we investigate the Ornstein–Uhlenback stochastic phase. We analyze the performance of this method for three key estimation problems: prediction, tracking, and smoothing. The results show significant reduction of the mean square error compared with the Mach–Zehnder interferometer under the same photon number flux inside the interferometers. In particular, we show that the method with the SU(1,1) interferometer can achieve fundamental quantum scaling, achieve stochastic Heisenberg scaling, and surpass the precision of the canonical measurement. © 2020 Chinese Laser Press

<https://doi.org/10.1364/PRJ.395682>

1. INTRODUCTION

Quantum optical phase estimation is a critical task in many applications such as quantum imaging [1–3], quantum sensing [4–7], and gravitational wave detection [8,9]. To date, most works have focused on the estimation of a constant phase φ , in which a Mach–Zehnder interferometer (MZI) is the most commonly used device [10,11]. The precision of estimation is limited by the shot noise when the classic resources are used. This limit is often called the standard quantum limit (SQL), $\Delta\varphi \propto 1/\sqrt{N}$, where N is the average number of photons in the probe state [12,13]. Many efforts have been taken to improve the precision. Most of them focus on utilizing nonclassical states to reduce the quantum noise, such as the squeezed states and entanglement states [14–16]. It has been shown that the maximally entangled number state ($N00N$) is the optimum probe state to reach the Heisenberg limit (HL) $\Delta\varphi \propto 1/N$ [17,18]. Moreover, for the constant phase estimation, the variance of estimation ($1/\sqrt{vN}$ or $1/\sqrt{v}N$) will decrease indefinitely as the number of measurement v increases.

However, it is not enough to just estimate the constant phase because many signals of interest in the real world are time varying and stochastic [19–23]. Thus, how to estimate such a time-varying phase with high precision is of practical importance. We assume that $\varphi(t)$ is the phase to be estimated. It can be treated as a constant in $t_i \leq t < t_i + dt$ if the dt is small enough, i.e., $\varphi(t) = \varphi_i$, so the phase can be discretized as $(\varphi_0, \varphi_1, \dots, \varphi_i, \dots, \varphi_{n-1}, \varphi_n)$. There is a set of observations $(r_0, r_1, \dots, r_i, \dots, r_{n-1}, r_n)$ used to estimate the $\{\varphi_i\}$, where r_i is the observation at time i . Compared to the estimation of constant phase, $\langle(\varphi_i - \varphi_{i\pm s})^2\rangle$ between the phases at two different times ($i, i \pm s$) increases as s increases. Therefore, the correlation between $r_{i\pm s}$ and φ_i decreases as s increases, and the number of observations that can be used to improve precision is limited. So there is a limited precision for the stochastic phase estimation, even for infinite measurement time. The mean square error (MSE) in the estimation of a stationary Gaussian stochastic phase with a power-law spectrum $\kappa^{p-1}/(\omega^p + \lambda^p)$ using coherent states scales as $(\kappa/N)^{(p-1)/p}$, which is called stochastic SQL.

Here \mathcal{N} is the photon flux [24–26]. Similar to constant phase estimation, there is a stochastic Heisenberg scaling for the estimation of this stochastic phase, which scales as $(\kappa/\mathcal{N})^{2(p-1)/(p+1)}$ [24–26]. Previous works show that the stochastic SQL can be surpassed with nonclassical resources such as the squeezed light with an adaptive quantum smoothing technique or canonical phase measurement [27–32]. However, practical methods to achieve quantum enhancement in the stochastic phase estimation still remain largely unexplored.

In this paper, we propose a method to estimate the stochastic phase using a new measurement device, the SU(1,1) interferometer. Such device, also known as the nonlinear interferometer (NLI), was originally proposed by Yurke and experimentally demonstrated with a signal-to-noise ratio (SNR) surpassing that of the MZI [33–43]. By combining the NLI with adaptive feedback technique, we show that MSE of the estimation is reduced over a range of parametric amplifier gain G compared to that of MZI. For a fixed photon flux, there is an optimal G that minimizes the MSE and maximizes the precision. In particular, with the optimal G , the precision of our scheme surpasses the stochastic SQL and achieves stochastic Heisenberg scaling asymptotically.

2. STOCHASTIC PHASE ESTIMATION SCHEME

The schematic diagram of the estimation of a stochastic phase with NLI is shown in Fig. 1, in which the NLI contains two parametric amplifiers (PAs). Two input modes of the first PA are injected with a coherent state $|\alpha\rangle$ and a vacuum state. The phase $\varphi(t)$ to be estimated is imposed to one arm of the interferometer. One of the output modes is measured with the homodyne measurement. The measured results after a displaced operation yield photocurrent $r(t)$. The phase $\Phi(t)$ in the other arm and the phase $\theta(t)$ of the local oscillator are adaptively controlled based on $\varphi_f(t)$, which is estimated from $r(s)$ for all the regions $s < t$. In the NLI, the first PA plays the role of beam splitting. If we define $\hat{c}_{in}, \hat{d}_{in}$ to be the annihilation

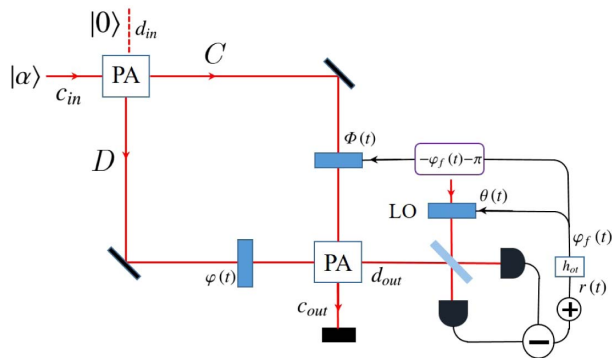


Fig. 1. Schematic diagram of enhanced stochastic phase estimation with an SU(1,1) interferometer. This interferometer consists of two parametric amplifiers (PAs), and the input states are the coherent state and the vacuum state. $\varphi(t)$ is the stochastic phase to be estimated, and the phase $\Phi(t)$ in the other arm is adaptively controlled. $r(t)$ is photocurrent, which is equal to the homodyne measurement results after an added operation. The phase $\theta(t)$ of the local oscillator is adaptively controlled simultaneously, and h_o is the optimum linear processor of phase tracking.

operators of the two inputs and \hat{C}, \hat{D} to be the annihilation operators of the outputs, the relation of the input/output of the PA can be written as $\hat{C} = G\hat{c}_{in} + g\hat{d}_{in}^\dagger$, $\hat{D} = G\hat{d}_{in} + g\hat{c}_{in}^\dagger$, where the gain G of the PA can also be parametrized by hyperbolic functions [44,45] and $G^2 - g^2 = 1$ [36]. The second PA, which has the same gain G , plays the role of recombination, so the complete input/output relation of the NLI is

$$\begin{aligned} \hat{c}_{out} &= G(G\hat{c}_{in} + g\hat{d}_{in}^\dagger)e^{i\Phi(t)} + g(g\hat{c}_{in} + G\hat{d}_{in}^\dagger)e^{-i\varphi(t)}, \\ \hat{d}_{out} &= g(G\hat{c}_{in}^\dagger + g\hat{d}_{in})e^{-i\Phi(t)} + G(g\hat{c}_{in}^\dagger + G\hat{d}_{in})e^{i\varphi(t)}. \end{aligned} \quad (1)$$

When we perform a homodyne detection at the output mode d_{out} , and the homodyne detection result is added by $2Gg|\alpha|\varphi_f(t)$, the photocurrent can be approximately represented as (Appendix A)

$$r(t) \approx \frac{2Gg|\beta|}{\sqrt{G^2 + g^2}}\varphi(t) + \sqrt{2G^2g^2\sigma_f^2 + 1}n(t). \quad (2)$$

Here we have adaptively controlled the feedback phase and the phase of the local oscillator to be $\Phi(t) = -\varphi_f(t) - \pi$, $\theta(t) = \varphi_f(t) + \pi/2$. This adjustment makes each measurement most sensitive and maximizes the phase information obtained. $n(t)$ is the normalized Gaussian white noise from the homodyne measurement, which satisfies $\langle n(t)n(s) \rangle = \delta(s - t)$. Moreover, the photocurrent has Gaussian stationary statistics, and $\sigma_f^2 = \langle [\varphi(t) - \varphi_f(t)]^2 \rangle$ is stationary MSE. We have defined the photon flux in the interferometer to be $|\beta|^2 = (G^2 + g^2)|\alpha|^2$.

According to the time span of the observations in the estimation, the time-varying phase estimation can be divided into three cases. Prediction: the future phase φ_{i+m} is estimated with observations (r_0, r_1, \dots, r_i) . Tracking: the current and previous observations (r_0, r_1, \dots, r_i) are used to estimate the current phase φ_i . Smoothing: the measurement results beyond the time i are also used to estimate phase φ_i . To analyze these three estimation problems simultaneously, we introduce a general desired signal $d(t) = \varphi(t + \varepsilon)$, which is estimated with photocurrent $r(s \leq t)$. Here ε can be any real number, and the three kinds of phase estimation can be defined according to the value of ε . Based on the measurement photocurrent $r(t)$, the desired signal can be estimated as

$$d_f(t) = \int_{-\infty}^t d\tau h_o(t, \tau)r(\tau), \quad (3)$$

where $h_o(t, \tau)$ is the impulse response function, which represents the output at time t if the input at time τ is an impulse, and it can be marked as $h_{op}(t, \tau)$, $h_{ot}(t, \tau)$, and $h_{os}(t, \tau)$ for prediction, tracking, and smoothing, respectively. The optimum linear processor $h_o(t, \tau)$ for minimizing the MSE $\xi(t) = \langle [d(t) - d_f(t)]^2 \rangle$ satisfies [46]

$$K_{dr}(t - \eta) = \int_{-\infty}^t h_o(t - \varepsilon)K_r(\varepsilon - \eta)d\varepsilon, \quad (4)$$

which is the Wiener–Hopf equation and $K_{dr}(t - \eta) = \langle d(t)r(\eta) \rangle$, $K_r(\varepsilon - \eta) = \langle r(\varepsilon)r(\eta) \rangle$. The correlation function only depends on the time difference due to the fact that received photocurrent and desired signal are jointly stationary and time invariant. We can obtain the optimum linear impulse response function h_o by solving the Wiener–Hopf equation,

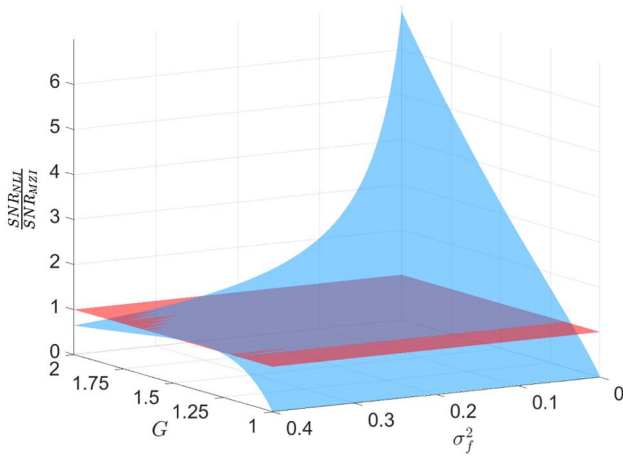


Fig. 2. Ratio of the two SNRs. The blue surface represents the ratio of the two SNRs. The red surface represents the case in which the two interferometers have equal SNR.

and the MSE of the phase estimation can be calculated at the same time.

In the MZI case, the splitting and recombination of light are accomplished by 50:50 beam splitters (BSs). The relation of the input/output of the beam splitter is $\hat{C} = 1/\sqrt{2}(\hat{c}_{in} + i\hat{d}_{in})$ and $\hat{D} = 1/\sqrt{2}(i\hat{c}_{in} + \hat{d}_{in})$. Similar to NLI, we set the two feedback phases $\Phi(t) = \varphi_f(t)$ and $\theta(t) = \varphi_f(t) + \pi$ for the most sensitive estimation. The photocurrent can be calculated as $r(t) \approx |\beta|\varphi(t) + n(t)$ (Appendix A). Here we should note that the two input modes of the first BS are injected with a coherent state $|\beta\rangle$ and a vacuum state, which makes the photon number flux inside both interferometers the same. From the two photocurrents, we can derive the relation $\frac{SNR_{NLI}}{SNR_{MZI}} = \frac{4G^2(G^2-1)}{(2G^2-1)[2G^2(G^2-1)\sigma_f^2+1]}$, where SNR_{NLI} and SNR_{MZI} are the SNRs of these two interferometers, respectively. Figure 2 shows the ratio of the two SNRs (blue surface). As G increases from 1, the SNR of the NLI increases with G , which agrees with the previous analysis of NLI [36]. Moreover, the two parameters (G, g) in the relation of the input/output of the PA are asymmetric. If the gain G is close to 1, for the fixed photon flux $|\beta|^2$, the part of signal of Eq. (2) is close to zero and the SNR of the NLI is smaller than that of the MZI. When G exceeds a certain threshold, the SNR of the NLI surpasses that of the MZI. For finite σ_f^2 , the further increase of G will reduce the SNR. This result can be understood from Eq. (2): when G is large, the signal term increases linearly with G while the noise term increases quadratically with G . Since the MSE in the estimation of time-varying phase cannot be arbitrarily small, we expect there is an optimal G for stochastic phase estimation, which is different from the case of measuring a constant phase.

3. ORNSTEIN-UHLENBACK STOCHASTIC PHASE ESTIMATION

As an example, we consider the situation that the time-varying phase $\varphi(t)$ to be estimated follows an Ornstein-Uhlenback

stochastic process, which can be found in many practical physical processes and is defined by [31]

$$\frac{d\varphi(t)}{dt} = -\lambda\varphi(t) + \sqrt{\kappa}\frac{dV(t)}{dt}. \quad (5)$$

Here λ^{-1} is the correlation time of $\varphi(t)$. $dV(t)$ represents the Wiener process, which satisfies $\langle dV(t)dV(s) \rangle = \delta(s-t)dt$. κ is the magnitude of the Wiener noise. The expectation value of $\varphi(t)$ is 0 and its statistics are stationary, which means the correlation between the phases at two different times only depends on their time difference. Moreover, the spectral density spectrum of $\varphi(t)$ is $S_\varphi(\omega) = \kappa/(\omega^2 + \lambda^2)$. In this situation, Eq. (4) can be solved with the Wiener technique, and the Fourier transform of the optimum linear response function h_o is (Appendix B)

$$H_o(\omega) = \begin{cases} \frac{\kappa\sqrt{P}e^{i\omega\varepsilon}}{N\lambda(1+\sqrt{1+\Lambda})(\lambda\sqrt{1+\Lambda}+i\omega)}, & \varepsilon > 0, \\ \frac{\kappa\sqrt{P}e^{i\omega\varepsilon}}{N[\lambda^2(1+\Lambda)+\omega^2]} \left[1 - \frac{e^{\varepsilon(\lambda\sqrt{1+\Lambda}-i\omega)}(\lambda+i\omega)}{\lambda(1+\sqrt{1+\Lambda})} \right], & \varepsilon \leq 0, \end{cases} \quad (6)$$

where $P = 4G^2g^2|\alpha|^2$, $\Lambda = P\kappa/N\lambda^2$, and $N = 2G^2g^2\sigma_f^2 + 1$. Equation (6) shows that the optimum linear processor is a low-pass filter with a cutoff frequency of $\lambda\sqrt{1+\Lambda}$ in the case of $\varepsilon = 0$. When this optimum linear processor is used to estimate the phase, the minimum MSE of the estimation is (Appendix C)

$$\xi = \langle [d(t) - d_f(t)]^2 \rangle = K_d(0) - \int_0^\infty K_{d_z}^2(\tau) d\tau, \quad (7)$$

where

$$K_{d_z}(\tau) = \begin{cases} \frac{\sqrt{P}\kappa}{\sqrt{N}\lambda} \frac{1}{1+\sqrt{1+\Lambda}} e^{-\lambda(\tau+\varepsilon)}, & \tau + \varepsilon \geq 0, \\ \frac{\sqrt{P}\kappa}{\sqrt{N}\lambda} \frac{1}{1+\sqrt{1+\Lambda}} e^{\lambda\sqrt{1+\Lambda}(\tau+\varepsilon)}, & \tau + \varepsilon < 0. \end{cases} \quad (8)$$

The stochastic phase is stationary and the optimum filter is time invariant, so the minimum MSE is a constant. To calculate the MSE analytically, we divide the discussion into three cases according to the value of ε : (i) $\varepsilon = 0$, (ii) $\varepsilon > 0$, and (iii) $\varepsilon < 0$. When $\varepsilon = 0$, $d(t) = \varphi(t)$. This is the phase-tracking case, and $\xi = \sigma_f^2$. The integral result ξ of Eq. (7) is still implicit because Λ is a function of σ_f^2 . After solving the implicit result, the minimum MSE of tracking is (Appendix C)

$$\sigma_f^2 = \frac{-\lambda - G^2g^2\kappa + \sqrt{(\lambda - G^2g^2\kappa)^2 + 4G^2g^2\left(\frac{|\beta|^2}{G^2+g^2} + \lambda\right)\kappa}}{4G^2g^2\left(\frac{|\beta|^2}{G^2+g^2} + \lambda\right)}. \quad (9)$$

Similarly, the MSEs of the other two cases can be calculated as (Appendix C)

$$\xi_{NLI} = \begin{cases} \frac{\kappa}{2\lambda} \left[1 - \frac{\Lambda}{(1+\sqrt{1+\Lambda})^2} e^{-2\lambda\varepsilon} \right], & \varepsilon > 0, \\ \frac{\kappa}{2\lambda} \left[\frac{1}{\sqrt{1+\Lambda}} + \frac{\Lambda e^{2\lambda\sqrt{1+\Lambda}\varepsilon}}{(1+\sqrt{1+\Lambda})^2\sqrt{1+\Lambda}} \right], & \varepsilon < 0, \end{cases} \quad (10)$$

where $\varepsilon > 0$ stands for the prediction of the future phase with current measurement outcomes and $\varepsilon < 0$ is the case of smoothing. Before investigating the enhancement of phase estimation with NLI, we set the precision of the

Ornstein–Uhlenback stochastic phase estimation with a coherent state and the MZI as the classical limit. To compare the MSEs of the two types of interferometers, we make the photon number flux inside them equivalent, i.e., $N_{\text{MZI}} = N_{\text{NLI}} = |\beta|^2$. In this case the MSE of phase estimation with the MZI can be written as (Appendix C)

$$\xi_{\text{MZI}} = \begin{cases} \frac{\kappa}{2\lambda} \left[1 - \frac{\Lambda_1}{(1+\sqrt{1+\Lambda_1})^2} e^{-2\lambda\varepsilon} \right], & \varepsilon > 0, \\ \frac{\kappa}{2\lambda} \left[\frac{1}{\sqrt{1+\Lambda_1}} + \frac{\Lambda_1 e^{2\lambda\sqrt{1+\Lambda_1}\varepsilon}}{(1+\sqrt{1+\Lambda_1})^2 \sqrt{1+\Lambda_1}} \right], & \varepsilon \leq 0, \end{cases} \quad (11)$$

where $\Lambda_1 = |\beta|^2 \kappa / \lambda^2$. To investigate the effect of the gain G on the MSE σ_f^2 , we consider the stochastic phase tracking with a fixed photon number flux $|\beta|^2 = 1.0 \times 10^7 \text{ s}^{-1}$, $\kappa = 1.0 \times 10^4 \text{ rad/s}$, and $\lambda = 1.0 \times 10^5 \text{ rad/s}$, and we vary G^2 from 1.1 to 50. Figure 3 shows that the MSE σ_f^2 can be reduced using the NLI compared to the classical limit with MZI. There is an optimal $G^2 = 7.4$ that gives the minimum MSE. For the feasible gain ($G^2 = 2.28$) with current techniques [47], we can already obtain 40.4% improvement in the MSE of phase tracking. This is expected from the analysis of the SNR. Here the optimal degree of gain depends on the photon number flux $|\beta|^2$, κ , and λ .

Figure 4 shows the MSE ξ as a function of $\lambda\varepsilon$ according to Eqs. (9)–(11), where the horizontal axis is the ratio between ε and the correlation time of $\varphi(t)$. Here we have set the parameters $\kappa = 1.0 \times 10^4 \text{ rad/s}$, $\lambda = 1.0 \times 10^5 \text{ rad/s}$, $G^2 = 7.4$, and $|\beta|^2 = 1.0 \times 10^7 \text{ s}^{-1}$. We can conclude three key implications from Fig. 4. First, the MSE becomes large with the increase of ε , and the smallest error is achieved with the smoothing. When ε is close to the correlation time, the MSEs tend to be the mean square variation of the stochastic phase $\kappa/2\lambda$, i.e., we cannot predict the phase away from coherence time. Second, the MSE of phase estimation with tracking is nearly two times of smoothing for both kinds of interferometers. Third, the MSEs of all cases are reduced significantly

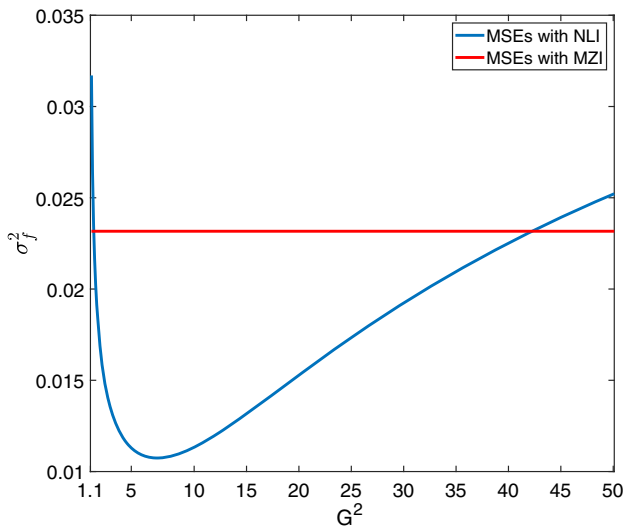


Fig. 3. MSE σ_f^2 of tracking as a function of G^2 for MZI (red line) and NLI (blue line). Here we take $\kappa = 1.0 \times 10^4 \text{ rad/s}$, $\lambda = 1.0 \times 10^5 \text{ rad/s}$, and $|\beta|^2 = 1.0 \times 10^7 \text{ s}^{-1}$.

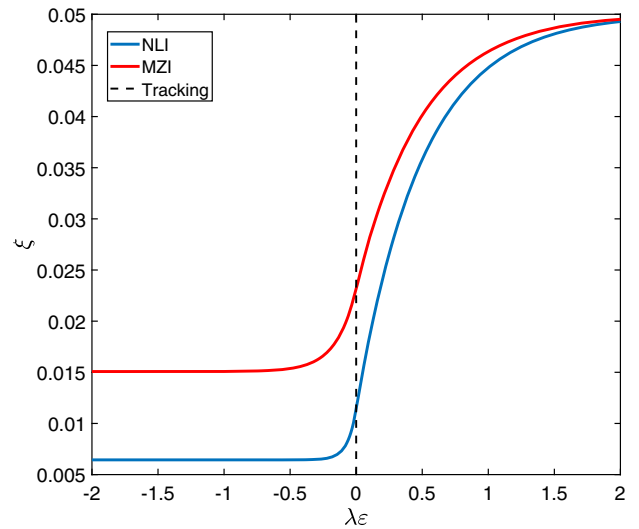


Fig. 4. MSE ξ as a function of $\lambda\varepsilon$ for MZI (red line) and NLI (blue line). The horizontal axis is the proportion between ε and the correlation time of $\varphi(t)$. The proportion equal to 0 represents phase tracking (black dotted line). $\lambda\varepsilon > 0$ and $\lambda\varepsilon < 0$ stand for prediction and smoothing, respectively. Here we make $\kappa = 1.0 \times 10^4 \text{ rad/s}$, $\lambda = 1.0 \times 10^5 \text{ rad/s}$, $G^2 = 7.4$, and $|\beta|^2 = 1.0 \times 10^7 \text{ s}^{-1}$.

below the classical limit (red line in Fig. 4) when we use the NLI.

So far we have shown that the stochastic phase estimation can be enhanced with NLI. In the following, we will demonstrate that the NLI achieves the Heisenberg scaling asymptotically. For a fixed photon flux $|\beta|^2$, the SNR of the measurement photocurrent Eq. (2) is $\text{SNR}_{\text{NLI}} = \frac{4G^2(G^2-1)|\beta|^2}{[(2G^4-G^2)\sigma_f^2+1](2G^2-1)}$. There is an optimal degree gain G_o maximizing the SNR, and the minimum MSE of phase tracking can be calculated as $\sigma_f^2 \approx 1/2G_o^4$. When $G_o^2 \gg 1$, $\frac{|\beta|^2}{2G_o^2} \gg \lambda$, $|\beta|^2 \gg \kappa$, and $G_o^4 \kappa \gg \lambda$, the optimal gain G_o meets the relationship $G_o^2 \approx (|\beta|^2 \kappa^2)^{1/3} / 2^{2/3} \kappa$ and we can obtain the tracking MSE (Appendix D)

$$\sigma_f^2 \approx 2^{1/3} \left(\frac{\kappa}{|\beta|^2} \right)^{2/3}. \quad (12)$$

Substituting this expression into Eq. (10) yields the MSE of smoothing

$$\xi \approx \left(\frac{\kappa}{2|\beta|^2} \right)^{2/3}, \quad (13)$$

which means the MSE of stochastic phase estimation with NLI can achieve the stochastic HL scaling [24–26]. Figure 5 shows the optimal smoothing MSE in the two kinds of interferometers and canonical measurement for different mean photon flux, which varies from $|\beta|^2 = 10^9 \text{ s}^{-1}$ to $|\beta|^2 = 10^{10} \text{ s}^{-1}$. It can be seen that the phase estimation with NLI has an enhancement on scaling compared with the classical limit $\xi \approx \frac{1}{2} \left(\frac{\kappa}{|\beta|^2} \right)^{1/2}$ using MZI, and the smoothing MSE can reach the stochastic Heisenberg scaling $O\left[\left(\frac{\kappa}{|\beta|^2}\right)^{2/3}\right]$ [25]. Moreover, using the NLI we can surpass the minimum MSE of canonical

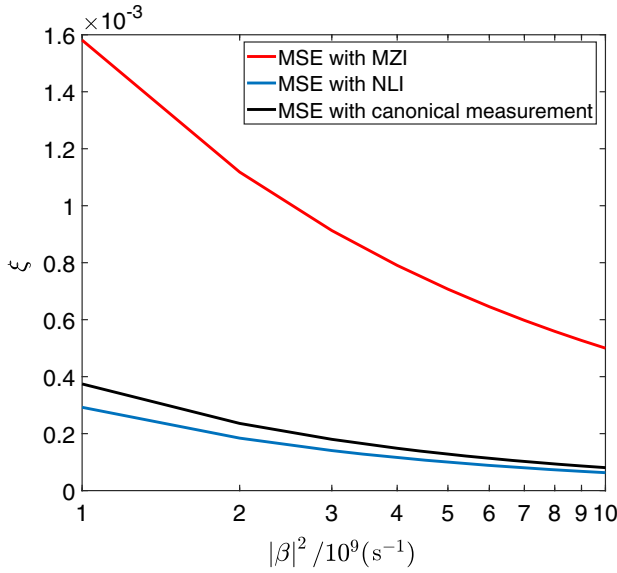


Fig. 5. Optimal smoothing MSE ξ as a function of photon number flux $|\beta|^2$ for MZI (red line), NLI (blue line), and canonical measurement (black line). Here we take $\kappa = 1.0 \times 10^4$ rad/s and $\lambda = 1.0 \times 10^5$ rad/s.

measurement, which is $\frac{4}{5}(\frac{\kappa}{|\beta|^2})^{2/3}$ [24]. It is worth noting that the MSE of tracking is nearly two times of smoothing for both interferometers. Therefore, the phase tracking can also reach the stochastic Heisenberg scaling. However, the MSE of stochastic phase prediction increases with the increase of ε , and we cannot predict the phase away from coherence time. So the prediction cannot reach the stochastic Heisenberg scaling when ε is relatively large.

4. CONCLUSION

In summary, we have proposed the stochastic optical phase estimation with an SU(1,1) interferometer. We find that a suitable range of parametric amplification gain can enhance the estimation, and there is an optimal gain minimizing the MSE, which is different from the estimation of constant phase [34,36]. Moreover, compared with the classical limit with MZI, the MSEs have significant reduction for prediction, tracking, and smoothing simultaneously under the same photon number flux inside the interferometers if we optimize the parametric amplifier gain. At last, we can achieve the stochastic Heisenberg scaling and surpass the minimum MSE by using the canonical measurement. These results highlight the advantages of the SU(1,1) interferometer in stochastic optical phase estimation and provide a new avenue for practical quantum metrology.

APPENDIX A. THE CALCULATION OF PHOTOCURRENT

When we perform homodyne detection at the output d_{out} , the measurement operator can be described as

$$\hat{X}_{\text{dout}}[\theta(t)] = \hat{d}_{\text{out}}^\dagger e^{i\theta(t)} + \hat{d}_{\text{out}} e^{-i\theta(t)}, \quad (\text{A1})$$

where $\theta(t)$ is the phase of the local oscillator. If two input modes of the first PA are injected with a coherent state

$|\alpha\rangle$ and a vacuum state, the mean value of homodyne measurement is

$$\begin{aligned} \langle \hat{X}_{\text{dout}}[\theta(t)] \rangle &= \langle \hat{d}_{\text{out}}^\dagger e^{i\theta(t)} + \hat{d}_{\text{out}} e^{-i\theta(t)} \rangle \\ &= 4Gg \cos \frac{\Phi(t) + \varphi(t)}{2} \\ &\quad \times \cos \left[\frac{\Phi(t) - \varphi(t)}{2} + \theta(t) \right] |\alpha| \\ &\approx 4Gg \sin \frac{\varphi(t) - \varphi_f(t)}{2} |\alpha| \\ &\approx 2Gg[\varphi(t) - \varphi_f(t)] |\alpha|, \end{aligned} \quad (\text{A2})$$

and the variance is

$$\begin{aligned} \Delta^2 \hat{X}_{\text{dout}}[\theta(t)] &= \langle \hat{X}_{\text{dout}}^2[\theta(t)] \rangle - \langle \hat{X}_{\text{dout}}[\theta(t)] \rangle^2 \\ &= 4G^2g^2\{1 + \cos[\Phi(t) + \varphi(t)]\} + 1 \\ &\approx 8G^2g^2 \left[\sin^2 \frac{\varphi(t) - \varphi_f(t)}{2} \right] + 1 \\ &\approx 2G^2g^2[\varphi(t) - \varphi_f(t)]^2 + 1. \end{aligned} \quad (\text{A3})$$

The photocurrent that we are interested in is

$$\begin{aligned} X_{\text{dout}}(t) &= \langle \hat{X}_{\text{dout}}[\theta(t)] \rangle + \Delta \hat{X}_{\text{dout}}[\theta(t)]n(t) \\ &\approx 2Gg[\varphi(t) - \varphi_f(t)] |\alpha| + \sqrt{2G^2g^2\sigma_f^2 + 1}n(t), \end{aligned} \quad (\text{A4})$$

where we have adaptively controlled the feedback phase in the other arm as $\Phi(t) = -\varphi_f(t) - \pi$, and the phase of the local oscillator was controlled as $\theta(t) = \varphi_f(t) + \pi/2$. $n(t)$ is a Gaussian white-noise term. Here the photocurrent has Gaussian stationary statistic, and $\sigma_f^2 = \langle [\varphi(t) - \varphi_f(t)]^2 \rangle_{ss}$ is stationary MSE. When the photocurrent is added by $2Gg\varphi_f(t)|\alpha|$, the photocurrent can be approximately represented as

$$\begin{aligned} r(t) &\approx 2Gg|\alpha|\varphi(t) + \sqrt{2G^2g^2\sigma_f^2 + 1}n(t) \\ &= \frac{2Gg|\beta|}{\sqrt{G^2 + g^2}}\varphi(t) + \sqrt{2G^2g^2\sigma_f^2 + 1}n(t), \end{aligned} \quad (\text{A5})$$

where we have defined that the photon flux inside the interferometer is $|\beta|^2 = (G^2 + g^2)|\alpha|^2$.

In the MZI case, the splitting and recombining of light are accomplished by 50:50 beam splitters. The relation of the input/output of the beam splitter is $\hat{C} = 1/\sqrt{2}(\hat{c}_{\text{in}} + i\hat{d}_{\text{in}})$ and $\hat{D} = 1/\sqrt{2}(i\hat{c}_{\text{in}} + \hat{d}_{\text{in}})$. The complete input/output relation of the MZI is

$$\begin{aligned} \hat{c}_{\text{out}} &= \frac{1}{2}\{[e^{i\varphi(t)} - e^{i\Phi(t)}]\hat{c}_{\text{in}} + i[e^{i\varphi(t)} + e^{i\Phi(t)}]\hat{d}_{\text{in}}\}, \\ \hat{d}_{\text{out}} &= \frac{1}{2}\{i[e^{i\varphi(t)} + e^{i\Phi(t)}]\hat{c}_{\text{in}} - [e^{i\varphi(t)} - e^{i\Phi(t)}]\hat{d}_{\text{in}}\}. \end{aligned} \quad (\text{A6})$$

Similar to the NLI, two input modes of the first BS are injected with a coherent state $|\beta\rangle$ and a vacuum state, and the mean value of homodyne measurement is

$$\begin{aligned}
& \langle \hat{X}_{\text{dout}}[\theta(t)] \rangle \\
&= \langle \hat{d}_{\text{out}}^{\dagger} e^{i\theta(t)} + \hat{d}_{\text{out}} e^{-i\theta(t)} \rangle \\
&= \left\{ 2 \cos \frac{\Phi(t) - \varphi(t)}{2} \cos \left[\frac{\Phi(t) + \varphi(t)}{2} - \theta(t) + \frac{\pi}{2} \right] \right\} |\beta| \\
&\approx 2|\beta| \cos \left[\frac{\varphi_f(t) + \varphi(t)}{2} - \theta(t) + \frac{\pi}{2} \right] \\
&\approx |\beta| [\varphi(t) - \varphi_f(t)], \tag{A7}
\end{aligned}$$

and the variance is

$$\Delta^2 \hat{X}_{\text{dout}}[\theta(t)] = \langle \hat{X}_{\text{dout}}^2[\theta(t)] \rangle - \langle \hat{X}_{\text{dout}}[\theta(t)] \rangle^2 \approx 1. \tag{A8}$$

The homodyne photocurrent is

$$\begin{aligned}
X_{\text{dout}}(t) &= \langle \hat{X}_{\text{dout}}[\theta(t)] \rangle + \Delta \hat{X}_{\text{dout}}[\theta(t)] n(t) \\
&\approx |\beta| [\varphi(t) - \varphi_f(t)] + n(t). \tag{A9}
\end{aligned}$$

Here the two feedback phases are $\Phi(t) = \varphi_f(t)$; $\theta(t) = \varphi_f(t) + \pi$. When the photocurrent is added by $|\beta|\varphi_f(t)$, the photocurrent can be approximately represented as

$$r(t) \approx |\beta|\varphi(t) + n(t). \tag{A10}$$

APPENDIX B. SOLUTION OF THE WIENER-HOPF EQUATION

When we set $\tau = t - \sigma$ and $v = t - \epsilon$, the Wiener-Hopf equation

$$K_{dr}(t - \sigma) = \int_{-\infty}^t h_o(t - \epsilon) K_r(\epsilon - \sigma) d\epsilon \tag{B1}$$

became

$$K_{dr}(\tau) = \int_0^{\infty} h_o(v) K_r(\tau - v) dv. \tag{B2}$$

Here we solve the equation with two steps. The first step of solving this equation is that we suppose there is a whitening filter impulse response $\omega(\tau, t)$, which can transfer $r(t)$ to white process $z(\tau)$, and the filtering process can be described as

$$z(\tau) = \int_{-\infty}^{\infty} r(t) \omega(\tau - t) dt. \tag{B3}$$

Taking the inverse Fourier transform on both sides, we can obtain

$$|W(\omega)|^2 S_r(\omega) = 1, \tag{B4}$$

where $W(\omega)$ is the transfer function of impulse response $\omega(\tau - t)$ and $S_r(\omega)$ is the spectrum density of $r(t)$. According to the Eq. (2) in the main text, we can calculate spectrum density as

$$S_r(\omega) = \frac{4G^2 g^2 |\alpha|^2 \kappa}{\omega^2 + \lambda^2} + 2G^2 g^2 \sigma_f^2 + 1. \tag{B5}$$

When we set $H^+(\omega) = \sqrt{N} \frac{i\omega + \lambda\sqrt{1+\Lambda}}{i\omega + \lambda}$, $\Lambda = \frac{P\kappa}{N\lambda^2}$, $N = 2G^2 g^2 \sigma_f^2 + 1$, and $P = 4G^2 g^2 |\alpha|^2$, the spectrum density can be decomposed as $S_r(\omega) = H^+(\omega)H(\omega)$. In this step, we can see that the transfer function that transfers $r(t)$ to white process $z(\tau)$ is $W(\omega) = \frac{1}{H^+(\omega)}$.

In the second step, we suppose that $f_o(t, \tau)$ is the impulse response of the optimum linear filter for estimating $d(t)$ with $z(\tau)$, so the corresponding Wiener-Hopf equation is

$$K_{dz}(\tau) = \int_0^{\infty} f_o(v) K_z(\tau - v) dv, \quad \tau \geq 0. \tag{B6}$$

Because $z(\tau)$ is a white process, therefore $f_o(\tau) = K_{dz}(\tau)$, and it can be found as

$$\begin{aligned}
K_{dz}(\tau) &= \left\langle d(t) \int_{-\infty}^{\infty} \omega(v) r(t - \tau - v) dv \right\rangle \\
&= \int_{-\infty}^{\infty} \omega(-\mu) K_{dr}(\tau - \mu) d\mu. \tag{B7}
\end{aligned}$$

If we take the inverse Fourier transform on both sides, we get

$$[S_{dz}(\omega)]_+ = [W^*(\omega)S_{dr}(\omega)]_+ = \left[\frac{S_{dr}(\omega)}{[H^+(\omega)]^*} \right]_+, \tag{B8}$$

where $[S_{dz}(\omega)]_+ = \int_0^{\infty} K_{dz}(\tau) e^{-j\omega\tau} d\tau$. Here we use the subscript $[\]_+$, which denotes that the integration time of the inverse transform from 0 to ∞ . In this step, we can see that the transfer function of optimum linear filter for estimating $d(t)$ with $z(\tau)$ is

$$F(\omega) = \left[\frac{S_{dr}(\omega)}{[H^+(\omega)]^*} \right]_+. \tag{B9}$$

After the two steps, we can see that the complete optimum linear processor in the frequency domain is

$$H_o(\omega) = \frac{F(\omega)}{H^+(\omega)} = \frac{1}{H^+(\omega)} \left[\frac{S_{dr}(\omega)}{[H^+(\omega)]^*} \right]_+. \tag{B10}$$

On the other hand, there is correlation $K_{dr}(\tau) = \langle d(t)r(t - \tau) \rangle = \sqrt{P}K_\varphi(\tau + \epsilon)$ and $S_{dr}(\omega) = \frac{\kappa\sqrt{P}e^{j\omega\epsilon}}{\omega^2 + \lambda^2}$. So

$$\begin{aligned}
S_{dz}(\omega) &= \frac{S_{dr}(\omega)}{[H^+(\omega)]^*} = \frac{\kappa\sqrt{P}e^{j\omega\epsilon}}{\omega^2 + \lambda^2} \frac{\lambda - i\omega}{\sqrt{N}(\lambda\sqrt{1+\Lambda} - i\omega)} \\
&= \frac{\kappa\sqrt{P}e^{j\omega\epsilon}}{\lambda + i\omega} \frac{1}{\sqrt{N}(\lambda\sqrt{1+\Lambda} - i\omega)} \\
&= \frac{\kappa\sqrt{P}e^{j\omega\epsilon}}{\sqrt{N}\lambda(1 + \sqrt{1+\Lambda})} \left(\frac{1}{\lambda + i\omega} + \frac{1}{\lambda\sqrt{1+\Lambda} - i\omega} \right). \tag{B11}
\end{aligned}$$

To find $[S_{dz}(\omega)]_+$, we take the inverse Fourier transform

$$\begin{aligned}
K_{dz}(\tau) &= \mathcal{F}^{-1}[S_{dz}(\omega)] \\
&= \mathcal{F}^{-1} \left[\frac{\kappa\sqrt{P}e^{j\omega\epsilon}}{\sqrt{N}\lambda(1 + \sqrt{1+\Lambda})} \right. \\
&\quad \left. \times \left(\frac{1}{\lambda + i\omega} + \frac{1}{\lambda\sqrt{1+\Lambda} - i\omega} \right) \right] \\
&= \frac{\kappa\sqrt{P}e^{-\lambda(\tau+\epsilon)}}{\sqrt{N}\lambda(1 + \sqrt{1+\Lambda})} u(\tau + \epsilon) \\
&\quad + \frac{\kappa\sqrt{P}e^{\lambda\sqrt{1+\Lambda}(\tau+\epsilon)}}{\sqrt{N}\lambda(1 + \sqrt{1+\Lambda})} u(-\tau - \epsilon), \tag{B12}
\end{aligned}$$

where $u(\tau)$ is the Heaviside function. When $\epsilon = 0$,

$$f_o(\tau) = K_{dz}(\tau) = \frac{\kappa\sqrt{P}e^{-\lambda\tau}}{\sqrt{N}\lambda(1 + \sqrt{1+\Lambda})} u(\tau) \tag{B13}$$

and

$$F(\omega) = [S_{dz}(\omega)]_+ = \frac{\kappa\sqrt{P}}{\sqrt{N}\lambda(1 + \sqrt{1 + \Lambda})} \frac{1}{\lambda + i\omega}, \quad (\text{B14})$$

so the complete optimum linear processor of phase tracking in the frequency domain is

$$\begin{aligned} H_{ot}(\omega) &= \frac{F(\omega)}{H^+(\omega)} \\ &= \frac{\kappa\sqrt{P}}{\sqrt{N}\lambda(1 + \sqrt{1 + \Lambda})} \frac{1}{\lambda + i\omega} \frac{\lambda + i\omega}{\sqrt{N}(\lambda\sqrt{1 + \Lambda} + i\omega)} \\ &= \frac{\kappa\sqrt{P}}{N\lambda(1 + \sqrt{1 + \Lambda})(\lambda\sqrt{1 + \Lambda} + i\omega)}. \end{aligned} \quad (\text{B15})$$

When $\varepsilon < 0$,

$$\begin{aligned} F(\omega) = [S_{dz}(\omega)]_+ &= \frac{\kappa\sqrt{P}}{\sqrt{N}} \left[\frac{e^{i\omega\varepsilon}}{(\lambda + i\omega)(\lambda\sqrt{1 + \Lambda} - i\omega)} \right. \\ &\quad \left. - \frac{e^{\varepsilon\lambda\sqrt{1 + \Lambda}}}{\lambda(1 + \sqrt{1 + \Lambda})(\lambda\sqrt{1 + \Lambda} - i\omega)} \right]. \end{aligned} \quad (\text{B16})$$

So the complete optimum linear processor of smoothing in the frequency domain is

$$\begin{aligned} H_{os}(\omega) = \frac{F(\omega)}{H^+(\omega)} &= \frac{\kappa\sqrt{P}}{\sqrt{N}} \left[\frac{e^{i\omega\varepsilon}}{(\lambda + i\omega)(\lambda\sqrt{1 + \Lambda} - i\omega)} \right. \\ &\quad \left. - \frac{e^{\varepsilon\lambda\sqrt{1 + \Lambda}}}{\lambda(1 + \sqrt{1 + \Lambda})(\lambda\sqrt{1 + \Lambda} - i\omega)} \right] \\ &\quad \times \frac{\lambda + i\omega}{\sqrt{N}(\lambda\sqrt{1 + \Lambda} + i\omega)} \\ &= \frac{\kappa\sqrt{P}e^{i\omega\varepsilon}}{N[\lambda^2(1 + \Lambda) + \omega^2]} \left[1 - \frac{e^{\varepsilon(\lambda\sqrt{1 + \Lambda} - i\omega)}(\lambda + i\omega)}{\lambda(1 + \sqrt{1 + \Lambda})} \right]. \end{aligned} \quad (\text{B17})$$

When $\varepsilon > 0$,

$$F(\omega) = [S_{dz}(\omega)]_+ = \frac{\kappa\sqrt{P}e^{i\omega\varepsilon}}{\sqrt{N}\lambda(1 + \sqrt{1 + \Lambda})} \frac{1}{\lambda + i\omega}, \quad (\text{B18})$$

so the complete optimum linear processor of prediction in the frequency domain is

$$\begin{aligned} H_{op}(\omega) &= \frac{F(\omega)}{H^+(\omega)} \\ &= \frac{\kappa\sqrt{P}e^{i\omega\varepsilon}}{\sqrt{N}\lambda(1 + \sqrt{1 + \Lambda})} \frac{1}{\lambda + i\omega} \frac{\lambda + i\omega}{\sqrt{N}(\lambda\sqrt{1 + \Lambda} + i\omega)} \\ &= \frac{\kappa\sqrt{P}e^{i\omega\varepsilon}}{N\lambda(1 + \sqrt{1 + \Lambda})(\lambda\sqrt{1 + \Lambda} + i\omega)}. \end{aligned} \quad (\text{B19})$$

Similarly, for the MZI case, the complete optimum linear processor in the frequency domain is

$$H_o(\omega) = \begin{cases} \frac{\kappa|\beta|e^{i\omega\varepsilon}}{\lambda(1 + \sqrt{1 + \Lambda_1})(\lambda\sqrt{1 + \Lambda_1} + i\omega)}, & \varepsilon > 0, \\ \frac{\kappa|\beta|e^{i\omega\varepsilon}}{\lambda^2(1 + \Lambda_1) + \omega^2} \left[1 - \frac{e^{\varepsilon(\lambda\sqrt{1 + \Lambda_1} - i\omega)}(\lambda + i\omega)}{\lambda(1 + \sqrt{1 + \Lambda_1})} \right], & \varepsilon \leq 0, \end{cases} \quad (\text{B20})$$

where we set $\Lambda_1 = \frac{|\beta|^2\kappa}{\lambda^2}$.

APPENDIX C. THE CALCULATION OF MINIMUM MEAN SQUARE ERROR

In this paper, the phase to be estimated is stationary and the optimum filter is time-invariant, so the MSE is time independent and can be calculated as

$$\begin{aligned} \xi(t) &= \left\langle \left[d(t) - \int_{-\infty}^t r(\tau)h_o(t - \tau)d\tau \right]^2 \right\rangle \\ &= K_d(0) - \int_{-\infty}^t h_o(t - \tau)K_{dr}(t - \tau)d\tau \\ &= K_d(0) - \int_0^\infty h_o(\gamma)K_{dr}(\gamma)d\gamma \\ &= K_d(0) - \int_0^\infty K_{dz}(t)d\tau \\ &\quad \times \left[\frac{1}{2\pi} \int_{-\infty}^\infty e^{-j\omega t} d\omega \frac{\int_{-\infty}^\infty K_{dr}(\tau)e^{j\omega\tau}d\tau}{H^+(\omega)} \right]. \end{aligned} \quad (\text{C1})$$

In the third line we let $t - \tau = \gamma$, and we substitute $h_o(\gamma)$ with the inverse transform of $H_o(\omega) = \frac{1}{H^+(\omega)} \int_0^\infty K_{dz}(t)e^{-j\omega t}dt$ in the last line. Moreover, from Eq. (B8) we can see $K_{dz}(t) = \frac{1}{2\pi} \int_{-\infty}^\infty e^{-j\omega t} d\omega \frac{1}{H^+(\omega)} \int_{-\infty}^\infty K_{dr}(\tau)e^{j\omega\tau}d\tau$. So the MSE can be expressed as

$$\xi(t) = K_d(0) - \int_0^\infty K_{dz}^2(t)dt, \quad (\text{C2})$$

where $K_d(0) = \frac{\kappa}{2\lambda}$, and K_{dz} can be found from Eq. (B12). When $\varepsilon = 0$, there is filtering with zero delay, which is the phase tracking case. The integral result is

$$\begin{aligned} \sigma_f^2 &= \frac{\kappa}{2\lambda} - \int_0^\infty \frac{P\kappa^2}{N\lambda^2(1 + \sqrt{1 + \Lambda})^2} e^{-2\lambda\tau}d\tau \\ &= \frac{\kappa}{2\lambda} \left[1 - \frac{\Lambda}{(1 + \sqrt{1 + \Lambda})^2} \right], \end{aligned} \quad (\text{C3})$$

which is still implicit because Λ is a function of σ_f^2 . After solving the implicit result, the MSE of phase tracking is

$$\sigma_f^2 = \frac{-(\lambda - G^2g^2\kappa) + \sqrt{(\lambda - G^2g^2\kappa)^2 + 4G^2g^2\left(\frac{|\beta|^2}{G^2 + g^2} + \lambda\right)\kappa}}{4G^2g^2\left(\frac{|\beta|^2}{G^2 + g^2} + \lambda\right)}. \quad (\text{C4})$$

When $\varepsilon > 0$, there is filtering with prediction, and the MSE of the prediction is

$$\begin{aligned} \xi_p &= \frac{\kappa}{2\lambda} - \int_0^\infty \frac{P\kappa^2}{N\lambda^2(1 + \sqrt{1 + \Lambda})^2} e^{-2\lambda(\tau + \varepsilon)}d\tau \\ &= \frac{\kappa}{2\lambda} \left[1 - \frac{\Lambda}{(1 + \sqrt{1 + \Lambda})^2} e^{-2\lambda\varepsilon} \right]. \end{aligned} \quad (\text{C5})$$

When $\varepsilon < 0$, it is the case of smoothing. The MSE of smoothing is

$$\begin{aligned}\xi_\varepsilon &= \frac{\kappa}{2\lambda} - \frac{P\kappa^2}{N\lambda^2} \frac{\int_0^{-\varepsilon} e^{2\lambda\sqrt{1+\Lambda}(\tau+\varepsilon)} d\tau + \int_{-\varepsilon}^{\infty} e^{-2\lambda(\tau+\varepsilon)} d\tau}{(1 + \sqrt{1+\Lambda})^2} \\ &= \frac{\kappa}{2\lambda} - \frac{P\kappa^2}{N\lambda^2} \frac{1}{(1 + \sqrt{1+\Lambda})^2} \left(\frac{1 - e^{2\lambda\sqrt{1+\Lambda}\varepsilon}}{2\lambda\sqrt{1+\Lambda}} + \frac{1}{2\lambda} \right) \\ &= \frac{\kappa}{2\lambda} \left[\frac{1}{\sqrt{1+\Lambda}} + \frac{\Lambda e^{2\lambda\sqrt{1+\Lambda}\varepsilon}}{(1 + \sqrt{1+\Lambda})^2 \sqrt{1+\Lambda}} \right].\end{aligned}\quad (C6)$$

For the MZI case, we set $H^+(\omega) = \sqrt{N} \frac{i\omega + \lambda \sqrt{1+\Lambda_1}}{i\omega + \lambda}$, $\Lambda_1 = \frac{P\kappa}{N\lambda^2}$, $N = 1$, $P = |\beta|^2$, and we use the same calculation method as for the NLI. The MSE of phase estimation with MZI can be written as

$$\xi_{\text{MZI}} = \begin{cases} \frac{\kappa}{2\lambda} \left[1 - \frac{\Lambda_1}{(1 + \sqrt{1+\Lambda_1})^2} e^{-2\lambda\varepsilon} \right], & \varepsilon > 0, \\ \frac{\kappa}{2\lambda} \left[\frac{1}{\sqrt{1+\Lambda_1}} + \frac{\Lambda_1 e^{2\lambda\sqrt{1+\Lambda_1}\varepsilon}}{(1 + \sqrt{1+\Lambda_1})^2 \sqrt{1+\Lambda_1}} \right], & \varepsilon \leq 0. \end{cases}\quad (C7)$$

APPENDIX D. THE STOCHASTIC HEISENBERG LIMIT WITH NLI

The signal-to-noise ratio of the photocurrent Eq. (2) in the main text is

$$\text{SNR}_{\text{NLI}} = \frac{4G^2 g^2 |\alpha|^2}{2G^2 g^2 \sigma_f^2 + 1} = \frac{4G^2 (G^2 - 1) |\beta|^2}{[(2G^4 - G^2) \sigma_f^2 + 1] (2G^2 - 1)}.\quad (D1)$$

The optimal gain G_o that minimizes the MSE of estimation is equal to that maximizes the signal-to-noise ratio. Taking the derivative of both sides of the equation with respect to G^2 , and letting the derivative be 0, we can obtain

$$\sigma_f^2 = \frac{4[2G_o^2(G_o^2 - 1) + 1]}{[4G_o^2(G_o^2 - 1)]^2} \approx \frac{1}{2G_o^4}.\quad (D2)$$

The approximate equals sign is true in the case of $G_o^2 \gg 1$. Then it is combined with Eq. (9) in the main text, and we can obtain

$$\begin{aligned}\frac{1}{2G_o^4} &= \frac{1}{8(G_o^4 - G_o^2) \left(\frac{|\beta|^2}{2G_o^2 - 1} + \lambda \right)} \left\{ -2[\lambda - (G_o^4 - G_o^2)\kappa] \right. \\ &\quad \left. + \sqrt{4[\lambda - (G_o^4 - G_o^2)\kappa]^2 + 16(G_o^4 - G_o^2) \left(\frac{|\beta|^2}{2G_o^2 - 1} + \lambda \right) \kappa} \right\}.\end{aligned}\quad (D3)$$

When $G_o^2 \gg 1$, $\frac{|\beta|^2}{2G_o^2} \gg \lambda$, $|\beta|^2 \gg \kappa$, and $G_o^4 \kappa \gg \lambda$, it can transform to

$$\frac{1}{2G_o^4} \approx \frac{-2[-(G_o^4)\kappa] + \sqrt{4[-(G_o^4)\kappa]^2 + 8G_o^2|\beta|^2\kappa}}{8(G_o^4) \left(\frac{|\beta|^2}{2G_o^2} \right)}.\quad (D4)$$

After the calculation, the optimal gain G_o is

$$G_o \approx \sqrt{\frac{(|\beta|^2 \kappa^2)^{1/3}}{2^{2/3} \kappa}}.\quad (D5)$$

Substituting it into Eq. (D2), we can obtain the tracking MSE

$$\sigma_f^2 \approx 2^{1/3} \left(\frac{\kappa}{|\beta|^2} \right)^{2/3}.\quad (D6)$$

Funding. National Key Research and Development Program of China (2017YFA0303703, 2019YFA0308704); National Natural Science Foundation of China (11664017, 11690032, 11874155, 61490711, 61975077, 91436211, 91836303); Fundamental Research Funds for the Central Universities (No. 020214380068); Natural Science Foundation of Shanghai (17ZR1442900); Nanjing University Innovation and Creative Program for PhD candidate (2016017); Basic Research Project of Shanghai Science and Technology Commission (20JC1416100).

Disclosures. The authors declare no conflicts of interest.

[†]These authors contributed equally to this work.

REFERENCES

1. M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, and W. P. Bowen, "Biological measurement beyond the quantum limit," *Nat. Photonics* **7**, 229–233 (2013).
2. T. Ono, R. Okamoto, and S. Takeuchi, "An entanglement-enhanced microscope," *Nat. Commun.* **4**, 2426 (2013).
3. G. Brida, M. Genovese, and I. R. Berchera, "Experimental realization of sub-shot-noise quantum imaging," *Nat. Photonics* **4**, 227–230 (2010).
4. S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, "Advances in photonic quantum sensing," *Nat. Photonics* **12**, 724–733 (2018).
5. C. Degen, F. Reinhard, and P. Cappellaro, "Quantum sensing," *Rev. Mod. Phys.* **89**, 035002 (2017).
6. C. Bonato, M. S. Blok, H. T. Dinani, D. W. Berry, M. L. Markham, D. J. Twitchen, and R. Hanson, "Optimized quantum sensing with a single electron spin using real-time adaptive measurements," *Nat. Nanotechnol.* **11**, 247–252 (2015).
7. G. M. D'Ariano and M. G. A. Paris, "Arbitrary precision in multipath interferometry," *Phys. Rev. A* **55**, 2267–2271 (1997).
8. R. X. Adhikari, "Gravitational radiation detection with laser interferometry," *Rev. Mod. Phys.* **86**, 121–151 (2014).
9. Y. Ma, H. Miao, B. H. Pang, M. Evans, C. Zhao, J. Harms, R. Schnabel, and Y. Chen, "Proposal for gravitational-wave detection beyond the standard quantum limit through EPR entanglement," *Nat. Phys.* **13**, 776–780 (2017).
10. L. Pezzé and A. Smerzi, "Mach-Zehnder interferometry at the Heisenberg limit with coherent and squeezed-vacuum light," *Phys. Rev. Lett.* **100**, 073601 (2008).
11. M. J. Holland and K. Burnett, "Interferometric detection of optical phase shifts at the Heisenberg limit," *Phys. Rev. Lett.* **71**, 1355–1358 (1993).
12. C. M. Caves, "Quantum-mechanical noise in an interferometer," *Phys. Rev. D* **23**, 1693–1708 (1981).
13. V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nat. Photonics* **5**, 222–229 (2011).
14. N. Thomas-Peter, B. J. Smith, A. Datta, L. Zhang, U. Dörner, and I. A. Walmsley, "Real-world quantum sensors: evaluating resources for precision measurement," *Phys. Rev. Lett.* **107**, 113603 (2011).
15. M. Xiao, L.-A. Wu, and H. J. Kimble, "Precision measurement beyond the shot-noise limit," *Phys. Rev. Lett.* **59**, 278–281 (1987).
16. S. Rosen, I. Afek, Y. Israel, O. Ambar, and Y. Silberberg, "Sub-Rayleigh lithography using high flux loss-resistant entangled states of light," *Phys. Rev. Lett.* **109**, 103602 (2012).
17. A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, "Quantum interferometric optical lithography: exploiting entanglement to beat the diffraction limit," *Phys. Rev. Lett.* **85**, 2733–2736 (2000).

18. Y. Israel, S. Rosen, and Y. Silberberg, "Supersensitive polarization microscopy using NOON states of light," *Phys. Rev. Lett.* **112**, 103604 (2014).
19. D. W. Berry and H. M. Wiseman, "Adaptive quantum measurements of a continuously varying phase," *Phys. Rev. A* **65**, 043803 (2002).
20. K. Iwasawa, K. Makino, H. Yonezawa, M. Tsang, A. Davidovic, E. Huntington, and A. Furusawa, "Quantum-limited mirror-motion estimation," *Phys. Rev. Lett.* **111**, 163602 (2013).
21. M. Tsang, H. M. Wiseman, and C. M. Caves, "Fundamental quantum limit to waveform estimation," *Phys. Rev. Lett.* **106**, 090401 (2011).
22. H. Miao, R. X. Adhikari, Y. Ma, B. Pang, and Y. Chen, "Towards the fundamental quantum limit of linear measurements of classical signals," *Phys. Rev. Lett.* **119**, 050801 (2017).
23. R. Jiménez-Martínez, J. Kołodziej, C. Troullinou, V. G. Lucivero, J. Kong, and M. W. Mitchell, "Signal tracking beyond the time resolution of an atomic sensor by Kalman filtering," *Phys. Rev. Lett.* **120**, 040503 (2018).
24. D. W. Berry, M. Tsang, M. J. Hall, and H. M. Wiseman, "Quantum Bell-Ziv-Zakai bounds and Heisenberg limits for waveform estimation," *Phys. Rev. X* **5**, 031018 (2015).
25. D. W. Berry, M. J. W. Hall, and H. M. Wiseman, "Stochastic Heisenberg limit: optimal estimation of a fluctuating phase," *Phys. Rev. Lett.* **111**, 113601 (2013).
26. H. T. Dinani and D. W. Berry, "Adaptive estimation of a time-varying phase with a power-law spectrum via continuous squeezed states," *Phys. Rev. A* **95**, 063821 (2017).
27. M. Tsang, "Optimal waveform estimation for classical and quantum systems via time-symmetric smoothing," *Phys. Rev. A* **80**, 033840 (2009).
28. M. Tsang, J. H. Shapiro, and S. Lloyd, "Quantum theory of optical temporal phase and instantaneous frequency. II. Continuous-time limit and state-variable approach to phase-locked loop design," *Phys. Rev. A* **79**, 053843 (2009).
29. M. Tsang, "Time-symmetric quantum theory of smoothing," *Phys. Rev. Lett.* **102**, 250403 (2009).
30. D. W. Berry and H. M. Wiseman, "Adaptive phase measurements for narrowband squeezed beams," *Phys. Rev. A* **73**, 063824 (2006).
31. T. A. Wheatley, D. W. Berry, H. Yonezawa, D. Nakane, H. Arai, D. T. Pope, T. C. Ralph, H. M. Wiseman, A. Furusawa, and E. H. Huntington, "Adaptive optical phase estimation using time-symmetric quantum smoothing," *Phys. Rev. Lett.* **104**, 093601 (2010).
32. H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arai, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph, H. M. Wiseman, E. H. Huntington, and A. Furusawa, "Quantum-enhanced optical-phase tracking," *Science* **337**, 1514–1517 (2012).
33. B. Yurke, S. L. McCall, and J. R. Klauder, "SU(2) and SU(1, 1) interferometers," *Phys. Rev. A* **33**, 4033–4054 (1986).
34. C. Brif and A. Mann, "Nonclassical interferometry with intelligent light," *Phys. Rev. A* **54**, 4505–4518 (1996).
35. J. Jing, C. Liu, Z. Zhou, Z. Y. Ou, and W. Zhang, "Realization of a nonlinear interferometer with parametric amplifiers," *Appl. Phys. Lett.* **99**, 011110 (2011).
36. Z. Y. Ou, "Enhancement of the phase-measurement sensitivity beyond the standard quantum limit by a nonlinear interferometer," *Phys. Rev. A* **85**, 023815 (2012).
37. F. Hudelist, J. Kong, C. Liu, J. Jing, Z. Ou, and W. Zhang, "Quantum metrology with parametric amplifier-based photon correlation interferometers," *Nat. Commun.* **5**, 3049 (2014).
38. B. E. Anderson, P. Gupta, B. L. Schmittberger, T. Horrom, C. Hermann-Avigliano, K. M. Jones, and P. D. Lett, "Phase sensing beyond the standard quantum limit with a variation on the SU(1, 1) interferometer," *Optica* **4**, 752–756 (2017).
39. M. Manceau, G. Leuchs, F. Khalili, and M. Chekhova, "Detection loss tolerant supersensitive phase measurement with an SU(1,1) interferometer," *Phys. Rev. Lett.* **119**, 223604 (2017).
40. S. Lemieux, M. Manceau, P. R. Sharapova, O. V. Tikhonova, R. W. Boyd, G. Leuchs, and M. V. Chekhova, "Engineering the frequency spectrum of bright squeezed vacuum via group velocity dispersion in an SU(1, 1) interferometer," *Phys. Rev. Lett.* **117**, 183601 (2016).
41. M. Gabrielli, L. Pezzè, and A. Smerzi, "Spin-mixing interferometry with Bose-Einstein condensates," *Phys. Rev. Lett.* **115**, 163002 (2015).
42. D. Linnemann, H. Strobel, W. Muessel, J. Schulz, R. Lewis-Swan, K. Kheruntsyan, and M. Oberthaler, "Quantum-enhanced sensing based on time reversal of nonlinear dynamics," *Phys. Rev. Lett.* **117**, 013001 (2016).
43. S. S. Szigeti, R. J. Lewis-Swan, and S. A. Haine, "Pumped-up SU(1,1) interferometry," *Phys. Rev. Lett.* **118**, 150401 (2017).
44. M. Ban, "Decomposition formulas for su(1, 1) and su(2) Lie algebras and their applications in quantum optics," *J. Opt. Soc. Am. B* **10**, 1347–1359 (1993).
45. G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Applications of the group SU(1,1) for quantum computation and tomography," *Laser Phys.* **16**, 1572–1581 (2006).
46. H. L. V. Trees and K. L. Bell, *Detection Estimation and Modulation Theory. Part I: Detection, Estimation, and Filtering Theory* (Wiley, 2013).
47. S. Liu, Y. Lou, J. Xin, and J. Jing, "Quantum enhancement of phase sensitivity for the bright-seeded SU(1,1) interferometer with direct intensity detection," *Phys. Rev. Appl.* **10**, 064046 (2018).