Improving the precision of weak measurements by postselection

Shengshi Pang and Todd A. Brun
Department of Electrical Engineering, University of Southern California, Los Angeles, California 90089, USA

Postselected weak measurement is a useful protocol to amplify weak physical effects. However, there has recently been controversy over whether it gives any advantage in precision. While it is now clear that retaining failed postselections can yield more Fisher information than discarding them, the advantage of postselecting weak measurement itself still remains to be clarified. In this paper, we address this problem by studying two widely used estimation strategies: averaging measurement results, and maximum likelihood estimation, respectively. For the first strategy, we find a surprising result that squeezed coherent states of the pointer can give postselected weak measurements a higher signal-to-noise ratio than non-postselected ones while all normal coherent states cannot, which suggests that raising the precision of weak measurements by postselection calls for the presence of “nonclassicality” in the pointer states. For the second strategy, we show that the quantum Fisher information of postselected weak measurements is generally larger than that of non-postselected weak measurements even without including the failed postselection events, but the gap can be closed with a proper choice of system states.

PACS numbers: 03.65.Ta, 03.65.Ca, 03.67.-a

Introduction.— Postselected weak measurement is a quantum measurement protocol first invented by Aharonov, Albert, and Vaidman in 1988 [1, 2]. It involves weak coupling between the system and the pointer, but the postselection on the system leads to a surprisingly counterintuitive effect: the average shift of the final pointer state can go far beyond the eigenvalue spectrum of the system observable, in sharp contrast to projective quantum measurements. The mechanism behind this effect is the coherence between the pointer states translated by different eigenvalues of the system observable. An enlightening interpretation of the large shift in a postselected weak measurement is based on the idea of superoscillation [3].

Postselected weak measurements have aroused enormous research interest in different fields, due to their potential application to observe tiny physical effects. Thanks to technical progress in recent years, weak measurements have been realized in experiments [4], and applied to measuring small parameters in various systems, including optical systems [5–14], SQUIDs [15, 16], and NMR [17]. More experimental protocols have also been proposed [18–20]. A general framework for weak measurement is given in [21], and reviews of the field are found in [22, 23].

One of the major goals in postselected weak measurements is to enhance the precision of estimating a small parameter in an interaction Hamiltonian. Starling et al.’s experiment [8] and Feizpour et al.’s proposal [20] showed that the signal-to-noise ratio (SNR) of weak measurements can be raised to the quantum limit by postselection, even in the presence of technical noise. Dressel et al. proposed a method of recycling failed postselection events to improve the precision [31]. However, recent intensive research has led to a negative conclusion [32–35, 37]. It was shown in [32–35] that the discarded failed postselections contain Fisher information, and the postselection probabilities themselves can also carry Fisher information [34], thus, filtering out the failed postselections will lead to a loss of precision. Nevertheless, it was later found [38] that the successful postselections can concentrate most of the Fisher information in the pointer, and the loss of Fisher information from discarding failed postselections can be reduced to the order of the small parameter [39], which is negligible for weak measurements. In addition, postselection-filtered weak measurement can also give better SNR in the presence of some specific technical noise [38].

As reviewed above, most of the previous research has been focused on whether failed postselections should be retained, provided the system is measured for postselection after the weak interaction with the pointer. However, a more fundamental problem is whether the system should be measured at all if one is to enhance the precision of weak measurements. If the measurements on the system could not give any advantage over (or even do worse than) non-postselected weak measurements, then it would become meaningless to study whether the failed postselections should be used or not. So, this question lies at the heart of postselected weak measurements: what is the significance of postselection measurements on the system in weak measurements?

At first glance, this question seems easy to answer, since postselection can amplify the signal, and thus the estimation SNR can be increased. However, the efficiency of postselection is rather low (roughly reciprocal to the amplification factor), which may cancel the benefits of the amplification effect in the SNR, so it is actually a very subtle problem. In fact, some known results have given negative answer to this question. For example, the numerical results in [40] showed that postselection on the system cannot improve the SNR with Gaussian pointer
states, and [37] further confirmed this observation (still with Gaussian pointer states) by analytical calculation of the Fisher information in the basis of coordinate space.

Nevertheless, a notable point is that the choice of system and pointer states were not optimized before, so it does not rule out the existence of other choices of the system and pointer states that may allow weak measurements with postselection to have higher precision. And the Gaussian states considered heretofore are quite “classical”, so it is of great interest whether using more “quantum” states can improve the precision of postselected weak measurements. Moreover, it is also possible that the Fisher information can be increased by using better measurements on the pointer, since the optimal measurements are not necessarily in the coordinate basis.

Answering these questions will clarify the advantages of postselection in weak measurements, and they are exactly the aim of this paper. We study the optimal precision of weak measurements for general system and pointer states, and compare postselected and non-postselected weak measurements. We consider two principal estimation strategies: one is averaging the measurement results of the pointer (AMR), and the other is maximum likelihood estimation (MLE). These two strategies have been widely used in practice.

For the strategy of AMR, an interesting result we find is that all normal (i.e., unsqueezed) coherent states do not give weak measurement an improvement in SNR with postselection, but properly squeezed coherent states do. This suggests that for the amplification of postselected weak measurements to enhance the precision, a necessary ingredient is some nonclassicality in the initial pointer states, which was missing from previous studies. This result greatly extends the understanding and feasibility of weak measurement in parameter estimation.

For the strategy of MLE, we study the quantum Fisher information, and show that even without considering the failed postselections, the quantum Fisher information of postselected weak measurements is generally higher than that of non-postselected weak measurements, and the increase of the quantum Fisher information does not vanish unless the initial state of the system becomes an eigenstate of the system observable in the interaction Hamiltonian.

**Weak value formalism.**— We start by reviewing the weak value formalism for postselected weak measurements. Suppose the initial state of the system is $|\Psi_i\rangle$ and the initial state of the pointer is $|D\rangle$. The interaction Hamiltonian between the system and the pointer is

$$H_{\text{int}} = gA \otimes \Omega \delta(t - t_0),$$

where the $\delta$ function indicates that the interaction is instantaneous at time $t_0$. Let $\hbar = 1$. After the interaction, the system is postselected to $|\Phi_f\rangle$, then the state of the pointer collapses to

$$|D_f\rangle = \frac{D_f}{\langle D_f|D_f\rangle} \langle \Phi_f| \exp(-igA \otimes \Omega) |\Phi_i\rangle |D\rangle \quad \text{(unnormalized)},$$

If one measures the change in an observable $M$ on the pointer, the average results is

$$\langle \Delta M \rangle_f = \frac{\langle D_f|M|D_f\rangle - \langle D|M|D\rangle}{\langle D_f|D_f\rangle}.$$ 

When $g$ is sufficiently small, $|D_f\rangle \approx \langle \Phi_f| \Phi_i\rangle (1 - igA_w \Omega) |D\rangle$, where $A_w$ is the weak value, defined as

$$A_w = \frac{\langle \Phi_f|A|\Phi_i\rangle}{\langle \Phi_f|\Phi_i\rangle}.$$  

The success probability of postselection is $P_s \approx |\langle \Phi_f|\Phi_i\rangle|^2$. It can be derived [38] that the average result of measuring $M$ on $|D_f\rangle$ is

$$\langle \Delta M \rangle_f \approx g \text{Im}A_w (\langle \{\Omega, M\}\rangle_D - 2\langle \Omega\rangle_D \langle M\rangle_D) + g \text{Re}A_w (\langle \Omega\rangle_D |\langle M\rangle_D| D),$$

where we have denoted $\langle D\rangle \cdot |D\rangle$ as $\langle \cdot |D\rangle$ for short.

The weak value [2] can be very large when $\langle \Phi_f|\Phi_i\rangle \ll 1$, and the dependence of $\langle \Delta M \rangle_f$ on $A_w$ in Eq. (3) indicates that the average shift can go beyond any eigenvalue of $A$ in this case. This is the origin of the weak measurement amplification.

**Optimal signal-to-noise ratio.**— We first study the precision of postselected weak measurements, and compare it with non-postselected weak measurements, to determine whether and when postselection can assist weak measurements in precision. To quantify the precision of estimating the parameter $g$, a widely used benchmark is the signal-to-noise ratio of the estimates, defined as

$$\text{SNR}_p = \frac{\sqrt{N} P_s (\Delta M)_f}{\sqrt{\text{Var}(M)_f}},$$

where $N$ is the total number of measurements. The factor $\sqrt{P_s}$ is because $\text{Var}(M)_f$ scales reciprocally to the number of successful postselections. In the first order approximation of $\text{SNR}_p$ with respect to $g$, the variance $\text{Var}(M)_D$ is equal to the variance of $M$ under the initial state of the pointer, i.e., $\text{Var}(M)_D = \text{Var}(M)|D\rangle$. With different pre- and postselections of the system, the $\text{SNR}_p$ is usually different, so a proper measure for the precisional performance of postselected weak measurements is the maximal SNR over all possible pre- and postselections. Direct maximization of $\text{SNR}_p$ by usual means (such as the variation method) is rather difficult, since the variation of $\text{SNR}_p$ gives a nonlinear equation that is not easy to deal with.

However, the results of [39] offer an alternative possible approach to this hard problem. In [39] it was shown that the largest success probability over all postselections of the system for a given weak value $A_w$ is

$$\max_{|\Psi_i\rangle} P_s \approx \frac{\text{Var}(A)_i}{\langle A_i^2\rangle_i - 2\langle A_i\rangle_i \text{Re}A_w + |A_w|^2},$$

where $|\cdot\rangle_i$ is short for $\langle \Phi_f_i| \Phi_i\rangle$. By exploiting this result, the task of maximizing the SNR over all pre- and postselections can be simplified to maximizing over all weak values $A_w$. 

Usually the weak value $A_w$ is complex, and can be denoted as $A_w = |A_w|e^{i\theta}$, so we can follow a two-step procedure to obtain the maxima of the SNR$_p$ over $A_w$: first, maximize SNR$_p$ over $|A_w|$, then maximize it over $\theta$. Below we show the outline of the derivation and the result of the optimal SNR, and leave the mathematical detail to [43].

It can be observed that an $|A_w|$ can be extracted from [3], then with [5], one can obtain [43] that SNR$_p$ reaches its maximum over $|A_w|$ when

$$|A_w| = \frac{(A^2)_i}{\langle A \rangle_i \cos \theta}. \quad (6)$$

Now, SNR$_p$ need only be maximized over $\theta$. To simplify this maximization, the dependence of $\text{SNR}_p$ on $\theta$ with $|A_w|$ given by (6) can be concentrated to the following factor

$$K(\theta) = \frac{\sin(\theta + \varphi)}{\sqrt{(A^2)_i - \langle A \rangle_i^2 \cos^2 \theta}}, \quad (7)$$

where

$$\varphi = \arctan \frac{i\langle \Omega, M \rangle_D}{\langle \Omega, M \rangle_D - 2\langle \Omega \rangle_D^2 (M)_D}. \quad (8)$$

So the problem boils down to maximizing $K(\theta)$ over $\theta$. In the supplemental material [43], the maximum $K(\theta)$ over $\theta$ is derived:

$$\max_{\theta} K(\theta) = \sqrt{\frac{\text{Var}(A)_i + \langle A \rangle_i^2 \sin^2 \varphi}{\text{Var}(A)_i \langle A^2 \rangle_i}}. \quad (9)$$

And the maximum of SNR$_p$ finally turns out to be

$$g(\varphi) = \frac{\sqrt{N} \langle \Omega, M \rangle_D - 2\langle \Omega \rangle_D^2 (M)_D^2 + \langle \Omega, M \rangle_D^2}}{\text{Var}(M)_D}. \quad (10)$$

where $\eta(\varphi) = \sqrt{\text{Var}(A)_i + \langle A \rangle_i^2 \sin^2 \varphi}$. This is the optimal SNR for postselected weak measurements. In [43], we also obtained an upper bound on SNR$_p$ based on (10).

When can postselection increase SNR? — The maximum SNR (10) quantifies the metrological performance of postselected weak measurements. To address the question that when (or whether) postselection can improve the SNR of weak measurements, we need to further compare (10) with the best SNR of non-postselected weak measurements.

In a weak measurement without postselection on the system, the average shift in the observable $M$ on the post-interaction pointer state is [43]

$$\langle \Delta M \rangle = i g\langle A \rangle_i \langle \Omega, M \rangle_D, \quad (11)$$

and $\max\langle A \rangle_i = \lambda_{\text{max}}(A)$, so the optimal SNR is

$$\max \text{SNR}_n = \frac{g \sqrt{N} \lambda_{\text{max}}(A) \langle \Omega, M \rangle_D}{\sqrt{\text{Var}(M)_D}}. \quad (12)$$

The ratio between the optimal SNR of postselected and non-postselected weak measurements is therefore

$$s = \frac{\eta(\varphi)}{\lambda_{\text{max}}(A)} \sqrt{1 + \frac{\langle \Omega, M \rangle_D^2 - 2 \langle \Omega \rangle_D^2 (M)_D^2}{\langle \Omega, M \rangle_D^2}}. \quad (13)$$

Using the definition of $\varphi$ from Eq. (8), one can show that $\langle \Omega, M \rangle_D^2 - 2\langle \Omega \rangle_D^2 (M)_D^2 = \text{csc}^2 \varphi$. So, $s$ can be reduced to

$$s = \frac{\sqrt{\text{Var}(A)_i \text{csc}^2 \varphi + \langle A \rangle_i^2}}{\lambda_{\text{max}}(A)} \quad (14)$$

Obviously, since $\text{csc}^2 \varphi \geq 1$, we have $\text{Var}(A)_i \text{csc}^2 \varphi + \langle A \rangle_i^2 \geq \langle A \rangle_i^2$. So when $|\Phi_i \rangle \rightarrow |\lambda_{\text{max}}(A)\rangle$, $\sqrt{\text{Var}(A)_i \text{csc}^2 \varphi + \langle A \rangle_i^2} \geq |\lambda_{\text{max}}(A)|$, and thus $s \geq 1$, which means that postselection in weak measurements will not reduce the SNR at least. But this is still not enough. The key question is: when (or whether) $\text{csc}^2 \varphi > 1$ can hold, so that postselection gives an increase of the SNR?

To answer this question, we move to the Fock space. Suppose that $\Omega = q, M = p$. Then $|\Omega, M \rangle_D = i$. In the Fock space, $q$ and $p$ can be represented by $q = a + a^\dagger)/\sqrt{2}, p = (a - a^\dagger)/\sqrt{2}$, so $\{q, p\} = i(a^2 - a^2)$, and $\text{csc}^2 \varphi = 1 + |\langle a^2 \rangle + a^2 - (a^2)^2|$. When the initial pointer state $|D\rangle$ is a normal coherent state, $\text{csc}^2 \varphi = 1$, so normal coherent states cannot give postselected weak measurements any advantage in SNR over non-postselected weak measurements. This suggests that “classical” pointer states are not able to improve the SNR of postselected weak measurements. It also generalizes the results of [37, 40].

An interesting question is whether “nonclassicality” in the pointer state can “activate” the advantage of postselected weak measurements in SNR.

Consider squeezed coherent states for the pointer. Suppose the initial state $|D\rangle$ of the pointer is

$$|\xi, \alpha\rangle = \exp \frac{1}{2} (\xi^* a^2 - \xi a^\dagger^2) |\alpha\rangle, \quad (15)$$

where $\xi$ is the squeeze parameter. Let $\xi = r e^{i\varphi}$, then it can be shown [43] that $\text{csc}^2 \varphi = 1 + 4(\sin \theta \sinh r \cosh r)^2$. Therefore

$$s = \frac{\sqrt{\langle A \rangle_i^2 + (1 + 4(\sin \theta \sinh r \cosh r)^2) \text{Var}(A)_i}}{\lambda_{\text{max}}(A)}. \quad (17)$$

It is clear from (17) that when $\sin \theta \neq 0$ and $\text{Var}(A)_i \neq 0$, one can acquire $s > 1$ with a large $r$, implying that the SNR of postselected weak measurements exceeds that of non-postselected weak measurements. So this shows that nonclassicality together with postselection can indeed improve the SNR of postselected weak measurements!
Figure 1. (Color online) The contour of $s$ is plotted for squeezed vacuum states $|\xi, 0\rangle = \exp \left(\xi^* a^2 - \xi a^2^2\right)|0\rangle$ with $|\xi| \leq 2$ and $|\arg \xi| \leq \pi$. The interaction Hamiltonian is $g r_{\xi} \otimes q$ with $g = 10^{-5}$. The weak value is fixed to 20. The momentum $p$ is measured on the pointer after the postselection. Each point in the figure represents a value of $s$, which is the ratio between the SNR of postselected and non-postselected weak measurements. It clearly shows that $|\xi|$ is large and $\arg \xi$ is not far away from $\pm 2\pi$, $|s|$ can be much large than 1, implying an increase in the SNR by postselecting the system.

To illustrate the above result, Fig. 1 plots the contour of the ratio $s$ on the complex plane for squeezed vacuum state $|\xi, 0\rangle$. Improvement of SNR can be explicitly observed in the figure.

**Optimal quantum Fisher information.**— Next, we turn to the SNR of weak measurements with maximum likelihood estimation. The exact SNR of MLE is usually difficult to obtain; however, a celebrated theorem by Fisher [44] shows that the Fisher information is an upper bound for the inverse variance of the MLE estimate, and it can be achieved in the asymptotic limit. So we will focus on the Fisher information for MLE here.

As different measurements on the pointer have different Fisher informations, a proper benchmark for the precision of MLE is the maximum Fisher information over all possible measurements (or, more generally, POVMs) on the pointer, which is defined as the quantum Fisher information [41, 42]. For a pure parameter-dependent state $|\psi_g\rangle$, the quantum Fisher information of estimating $g$ from many copies of this state is

$$ F^{(Q)} = 4(\langle \partial_g \psi_g | \partial_g \psi_g \rangle - |\langle \psi_g | \partial_g \psi_g \rangle|^2). $$

(18)

In postselected weak measurements, the pointer state after postselection on the system is $|D_f\rangle \approx e^{-i g A_w \Omega} |D\rangle$, so $|\partial_g D_f\rangle \approx -(i A_w \Omega + \langle \Omega \rangle_D |A_w\rangle |D\rangle$, and the quantum Fisher information is approximately [43]

$$ F_p^{(Q)} \approx 4 P_a |A_w|^2 \text{Var}(\Omega) D. $$

(19)

where we note the dependence on the postselection probability $P_a$. The maximum $P_a$ is given by $\min(|\langle \Omega \rangle_D|, 1)$, therefore, the maximum quantum Fisher information over all postselections given the weak value $A_w$ is

$$ F_p^{(Q)} \approx \frac{4|A_w|^2 \text{Var}(A_w) \text{Var}(\Omega) D}{|\langle A_w \rangle^2 - \langle A_w \rangle^2 \text{cos}^2 \theta|}. $$

(20)

Now, the task is to maximize $F_p^{(Q)}$ over $A_w$ to find the optimal quantum Fisher information. Suppose $A_w = |A_w|e^{i \theta}$, then the quantum Fisher information $F_p^{(Q)}$ can be maximized in a two-step procedure similar to how we maximized SNR. The critical $|A_w|$ is again given by $\text{Re} A_w + |A_w|^2$, leading to

$$ \max_{|A_w|} F_p^{(Q)} \approx \frac{4|A_w|^2 \text{Var}(A_w) \text{Var}(\Omega) D}{(\langle A_w \rangle^2 - \langle A_w \rangle^2 \text{cos}^2 \theta)}. $$

(21)

Obviously, when $\text{cos} \theta = \pm 1$, $F_p^{(Q)}$ is maximized, so the maximum $F_p^{(Q)}$ is

$$ \max_{A_w} F_p^{(Q)} \approx 4 \langle A_w \rangle^2 \text{Var}(\Omega) D. $$

(22)

As a comparison, consider non-postselected weak measurements. In a non-postselected weak measurement, the post-interaction pointer state is generally a mixed state since the pointer is entangled with the system after the weak interaction. The calculation of quantum Fisher information for mixed states is much more involved than for pure states, and an analytical result is generally unavailable. However, the weak measurement condition $g \ll 1$ can significantly reduce this difficulty, since the post-interaction pointer state is approximately $|D_f\rangle \approx e^{-ig(A_w \Omega)} |D\rangle$ when $g \ll 1$ [43].

By this simplification, we can immediately derive the quantum Fisher information for non-postselected weak measurements:

$$ F_n^{(Q)} \approx 4 \langle A_w \rangle^2 \text{Var}(\Omega) D. $$

(23)

Then, comparing $F_n^{(Q)}$ with $F_p^{(Q)}$, we can obtain the ratio between these two quantum Fisher information:

$$ \eta = \frac{F_p^{(Q)}}{F_n^{(Q)}} \approx \frac{\langle A_w \rangle^2}{\langle A_w \rangle^2}. $$

(24)

It is obvious that $\eta \geq 1$, implying that postselected weak measurements generally possess more quantum Fisher information than non-postselected weak measurements. However, when the initial state of the system approaches an eigenstate of $A$, the quantum Fisher information of non-postselected weak measurements can equal that of postselected weak measurements.

**Remark.**— It has been shown in previous research that the quantum Fisher information using failed postselections can have slightly more quantum Fisher information than without using them, up to the order of $g$ [38, 39]. If
we denote the quantum Fisher information of weak measurements using all failed postselections as \( F_{all}^{(Q)} \), then combining this with the above results produces

\[
F_{all}^{(Q)} \geq F_p^{(Q)} \geq F_n^{(Q)},
\]

which clearly shows the relation of the quantum Fisher information between the three types of weak measurements, and clarifies the metrological advantage of postselected weak measurements.

Acknowledgement.— This research was supported by the ARO MURI under Grant No. W911NF-11-1-0268.

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. 60, 1351 (1988).
[2] I. M. Duck, P. M. Stevenson and E. C. G. Sudarshan, Phys. Rev. D 40, 2112 (1989).
[3] M. V. Berry and P. Shukla, J. Phys. A: Math. Theor. 45, 015301 (2012).
[4] N. W. M. Ritchie, J. G. Story, and Randall G. Hulet, Phys. Rev. Lett. 88, 041803 (2009).
[5] O. Hosten and P. Kwiat, Science 319, 787 (2008).
[6] Y. Gorodetski, K. Y. Blökh, B. Stein, C. Genet, N. Shitrit, V. Kleiner, E. Hasman, and T. W. Ebbesen, Phys. Rev. Lett. 109, 013901 (2012).
[7] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Phys. Rev. Lett. 102, 173601 (2009).
[8] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A 80, 041803 (2009).
[9] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A 82, 063822 (2010).
[10] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, Phys. Rev. A 82, 011802(R) (2010).
[11] X.-Y. Xu, Y. Kedem, K. Sun, L. Vaidman, C.-F. Li, and G.-C. Guo, Phys. Rev. Lett. 111, 033604 (2013).
[12] P. Egan and J. A. Stone, Opt. Lett. 37, 4991 (2012).
[13] G. I. Viza, J. Martínez-Rincón, G. A. Howland, H. Frostig, I. Shomroni, B. Dayan, and J. C. Howell, Opt. Lett. 38, 2949 (2013).
[14] O. S. Magaia-Loaiza, M. Mirhosseini, B. Rodenburg, and R. W. Boyd, Phys. Rev. Lett. 112, 200401 (2014).
[15] P. Campagne-Ibarcq, L. Bretheau, E. Flurin, A. Auffèves, F. Mallet, and B. Huard, Phys. Rev. Lett. 112, 180402 (2014).
[16] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, Phys. Rev. Lett. 111, 090506 (2013).
[17] D. Lu, A. Brodutch, J. Li, H. Li, and R. Laflamme, New J. Phys. 16, 053015 (2014).
[18] A. Romito and Y. Gefen, Phys. Rev. Lett. 100, 056801 (2008).
[19] V. Shpitalnik, Y. Gefen and A. Romito, Phys. Rev. Lett. 101, 226802 (2008).
[20] N. Brunner and C. Simon, Phys. Rev. Lett. 105, 010405 (2010).
[21] O. Zilberberg, A. Romito, and Y. Gefen, Phys. Rev. Lett. 106, 080405 (2011).
[22] S. Wu and M. Zukowski, Phys. Rev. Lett. 108, 080403 (2012).
[23] G. Strübi and C. Bruder, Phys. Rev. Lett. 110, 083605 (2013).
[24] M. Bula, K. Bartkiewicz, A. Černoch, and K. Lemr, Phys. Rev. A 87, 033826 (2013).
[25] E. Meyer-Scott, M. Bula, K. Bartkiewicz, A. Černoch, J. Soubusta, T. Jennewein, and K. Lemr, Phys. Rev. A 88, 012327 (2013).
[26] J. Dressel, K. Lyons, A. N. Jordan, T. M. Graham, and P. G. Kwiat, Phys. Rev. A 88, 023821 (2013).
[27] S. Wu and Y. Li, Phys. Rev. A 83, 052106 (2011).
[28] A. G. Kofman, S. Ashhab, F. Nori, Phys. Rep. 520, 43 (2012).
[29] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, R. W. Boyd, Rev. Mod. Phys. 86, 307 (2014).
[30] A. Feizpour, X. Xing, and A. M. Steinberg, Phys. Rev. Lett. 107, 133603 (2011).
[31] J. Dressel, K. Lyons, A. N. Jordan, T. M. Graham, and P. G. Kwiat, Phys. Rev. A 88, 023821 (2013).
[32] G. C. Knee, G. A. D. Briggs, S. C. Benjamin, and E. M. Gauger, Phys. Rev. A 87, 012115 (2013).
[33] S. Tanaka and N. Yamamoto, Phys. Rev. A 88, 042116 (2013).
[34] L. Zhang, A. Datta, I. A. Walmsley, arXiv/quant-ph: 1310.5302 (2013).
[35] C. Ferrie and J. Combes, Phys. Rev. Lett. 112, 040406 (2014).
[36] L. Vaidman, arXiv/quant-ph: 1402.0199 (2014).
[37] G. C. Knee and E. M. Gauger, Phys. Rev. X 4, 011032 (2014).
[38] A. N. Jordan, J. Martínez-Rincón, and J. C. Howell, Phys. Rev. X 4, 011031 (2014).
[39] S. Pang, J. Dressel and T. A. Brun, Phys. Rev. Lett. 113, 030401 (2014).
[40] X. Zhu, Y. Zhang, S. Pang, C. Qiao, Q. Liu, and S. Wu, Phys. Rev. A 84, 052111 (2011).
[41] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. 72, 3439 (1994).
[42] S. L. Braunstein, C. M. Caves and G. J. Milburn, Ann. Phys. 247, 135 (1996).
[43] Supplemental material.
[44] R. A. Fisher, Proc. Camb. Soc. 22, 700 (1925).
SUPPLEMENTAL MATERIAL

I. AVERAGE SHIFT OF THE POINTER

In this section, we show how to derive the average shift of the pointer for postselected weak measurements and non-postselected weak measurements, respectively.

A. Postselected weak measurements

The interaction Hamiltonian between the system and the pointer is

$$H_{\text{int}} = gA \otimes \Omega \delta(t - t_0). \quad (S1)$$

Suppose the initial state of the system is $$|\Phi_i\rangle$$, and the initial state of the pointer is $$|D\rangle$$. The system is postselected in the state $$|\Phi_f\rangle$$ after the interaction $$(S1)$$, and the pointer collapses to the state (unnormalized)

$$|D_f\rangle = \langle \Phi_f | \exp(-igA \otimes \Omega) |\Phi_i\rangle |D\rangle. \quad (S2)$$

When $$g$$ is sufficiently small, $$|D_f\rangle$$ is approximately

$$|D_f\rangle \approx \langle \Phi_f | (1 - igA \otimes \Omega) |\Phi_i\rangle |D\rangle \approx \langle \Psi_f | \Psi_i \rangle (1 - igA_w \Omega) |D\rangle, \quad (S3)$$

where $$A_w$$ is the so-called weak value, defined as $$A_w = \frac{\langle \Psi_f | A |\Psi_i\rangle}{\langle \Psi_f | \Psi_i \rangle}$$. The success probability of postselection is

$$P_s \approx |\langle \Psi_f | \Psi_i \rangle|^2. \quad (S4)$$

The average shift in the measured value of the pointer observable $$M$$ is

$$\langle \Delta M \rangle_f = \langle D_f | M |D_f\rangle - \langle M \rangle_D. \quad (S5)$$

We denote $$\langle D | \cdot | D \rangle$$ as $$\langle \cdot \rangle_D$$ for short throughout the paper. Note that

$$\langle D_f | M |D_f\rangle \approx |\langle \Psi_f | \Psi_i \rangle|^2 \langle M \rangle_D + ig \text{Re} A_w \langle \Omega, M \rangle_D + g \text{Im} A_w \langle \{\Omega, M\} \rangle_D, \quad (S6)$$

$$\langle D_f | D_f\rangle \approx |\langle \Psi_f | \Psi_i \rangle|^2 (1 + 2g \text{Im} A_w \langle \Omega \rangle_D).$$

By plugging $$(S6)$$ into $$(S5)$$, we see that

$$\langle \Delta M \rangle_f = ig \text{Re} A_w \langle \{\Omega, M\} \rangle_D - 2\langle \Omega \rangle_D \langle M \rangle_D$$

$$+ ig \text{Im} A_w \langle [\Omega, M] \rangle_D, \quad (S7)$$

which is Eq. (3) in the main text.

B. Non-postselected weak measurements

In a non-postselected weak measurement, the interaction between the system and pointer is also given by $$(S1)$$, but there is no measurement on the system, so the post-interaction system-pointer state is $$\exp(-igA \otimes \Omega) |\Phi_i\rangle |D\rangle$$, and the average shift of the pointer is

$$\langle \Delta M \rangle_f = \langle \Phi_i | (D | \exp(igA \otimes \Omega) M \exp(-igA \otimes \Omega) |\Phi_i\rangle |D\rangle - \langle M \rangle_D. \quad (S8)$$

When $$g \ll 1$$,

$$\exp(igA \otimes \Omega) M \exp(-igA \otimes \Omega) \approx (I + igA \otimes \Omega) M (I - igA \otimes \Omega) \approx M + igA \otimes [\Omega, M]. \quad (S9)$$

Therefore,

$$\langle \Delta M \rangle_f = ig \langle A \rangle_i \langle [\Omega, M] \rangle_D. \quad (S10)$$
II. MAXIMIZATION OF SNR

In this section, we detail the optimization of the signal-to-noise ratio (SNR) over the weak value $A_w$ for postselected weak measurements by the two-step procedure outlined in the main text. We will also obtain an upper bound on the maximum SNR as a by-product.

The SNR of the estimated value of $g$ by a postselected weak measurement with interaction $\text{(S1)}$ is

$$ SNR_p = \frac{\sqrt{NP_s(\Delta M)_f}}{\sqrt{\text{Var}(M)_f}}. \quad (S11) $$

Since the numerator of $SNR_p$ is $O(g)$, we need only approximate the denominator of $SNR_p$ to $O(1)$ to guarantee that the $SNR_p$ has a precision of $O(g)$. Thus, we can assume $\text{Var}(M)_f \approx \text{Var}(M)_D$.

The postselection probability $P_s$ can take different values for a given weak value $A_w$ by varying the pre- and postselections. However, the maximium $P_s$ given the weak value $A_w$ is $[1]$

$$ \max P_s \approx \frac{\text{Var}(A)_{ii}}{\langle A^2 \rangle_{ii} - 2\langle A \rangle_{ii}\text{Re}A_w + |A_w|^2}. \quad (S12) $$

Thus, the $SNR_p$ can be written as

$$ SNR_p = g\sqrt{N} \frac{\text{Var}(A)_{ii}}{\text{Var}(M)_D} \frac{\text{Im}A_w(\langle{\Omega},M\rangle)_D - 2\langle{\Omega}\rangle_D\langle M \rangle_D + \text{i}\text{Re}A_w(\langle{\Omega},M\rangle)_D}{\sqrt{\langle A^2 \rangle_{ii} - 2\langle A \rangle_{ii}\text{Re}A_w + |A_w|^2}}. \quad (S13) $$

Let $A_w = |A_w|e^{i\theta}$. We can maximize $SNR_p$ over $A_w$ by a two-step procedure: first, maximize $SNR_p$ over $|A_w|$, then over $\theta$. We first maximize $SNR_p$ over $|A_w|$. Note that $\text{(S13)}$ can be written

$$ SNR_p = g\sqrt{N} \frac{\text{Var}(A)_{ii}}{\text{Var}(M)_D} \frac{\langle{\Omega},M\rangle_D - 2\langle{\Omega}\rangle_D\langle M \rangle_D + \text{Re}A_w(\langle{\Omega},M\rangle)_D}{\sqrt{\langle A^2 \rangle_{ii} - 2\langle A \rangle_{ii}\text{Re}A_w + |A_w|^2}} + i\text{Re}A_w(\langle{\Omega},M\rangle)_D \cos \theta \sqrt{\langle A^2 \rangle_{ii} - 2\langle A \rangle_{ii}\text{Re}A_w + |A_w|^2} + 1. \quad (S14) $$

The minimum of $\langle A^2 \rangle_{ii}|A_w|^{-2} - 2\langle A \rangle_{ii}\cos \theta|A_w|^{-1} + 1$ is

$$ 1 - \frac{\langle A \rangle_{ii}^2 \cos^2 \theta}{\langle A^2 \rangle_{ii}}, \quad (S15) $$

if $\cos \theta > 0$, and the critical point of $|A_w|$ is

$$ |A_w| = \frac{\langle A^2 \rangle_{ii}}{\langle A \rangle_{ii} \cos \theta}. \quad (S16) $$

So, the maximum of $SNR_p$ over $|A_w|$ is

$$ \max_{|A_w|} SNR_p = g\sqrt{N} \frac{\text{Var}(A)_{ii}}{\text{Var}(M)_D} \frac{(\langle{\Omega},M\rangle)_D - 2\langle{\Omega}\rangle_D\langle M \rangle_D + \text{i}\text{Re}A_w(\langle{\Omega},M\rangle)_D}{\sqrt{\langle A^2 \rangle_{ii} - \langle A \rangle_{ii}^2 \cos^2 \theta}} \cos \theta \sqrt{\langle A^2 \rangle_{ii} - 2\langle A \rangle_{ii}\text{Re}A_w + |A_w|^2} + 1. \quad (S17) $$

Next, we maximize $\text{(S17)}$ over $\theta$. For simplicity, let

$$ \varphi = \arctan \frac{i(\langle{\Omega},M\rangle)_D}{(\langle{\Omega},M\rangle)_D - 2\langle{\Omega}\rangle_D\langle M \rangle_D}. \quad (S18) $$

Then $\text{(S17)}$ can be simplified to

$$ \max_{|A_w|} SNR_p = gK(\theta)\sqrt{N} \frac{\text{Var}(A)_{ii}}{\text{Var}(M)_D} \sqrt{((\langle{\Omega},M\rangle)_D - 2\langle{\Omega}\rangle_D\langle M \rangle_D)^2 + (\langle{\Omega},M\rangle)_D^2)}, \quad (S19) $$

where

$$ K(\theta) = \frac{\sin(\theta + \varphi)}{\sqrt{\langle A^2 \rangle_{ii} - \langle A \rangle_{ii}^2 \cos^2 \theta}}. \quad (S20) $$
The key is to maximize $K(\theta)$ over $\theta$ to obtain the maximum SNR. Note that $K(\theta)$ can be rewritten as

$$K(\theta) = \frac{\sin(\theta + \varphi)}{\sqrt{\text{Var}(A)_i + (\langle A \rangle^2_i \sin(\theta + \varphi) \cos(\theta + \varphi) \sin \varphi)^2}}$$

$$= \frac{1}{\sqrt{\text{Var}(A)_i + (\langle A \rangle^2_i \cos^2 \varphi) - 2\langle A \rangle^2_i \sin \varphi \cos \varphi \cot(\theta + \varphi) + (\text{Var}(A)_i + \langle A \rangle^2_i \sin^2 \varphi) \cot^2(\theta + \varphi)}}.$$  

(S21)

It is obvious that $K(\theta)$ is maximized when

$$\cot(\theta + \varphi) = \frac{\langle A \rangle^2_i \sin \varphi \cos \varphi}{\text{Var}(A)_i + \langle A \rangle^2_i \sin^2 \varphi},$$

(S22)

and the maximum of $K(\theta)$ is

$$\max_{\theta} K(\theta) = \frac{1}{\sqrt{\text{Var}(A)_i + (\langle A \rangle^2_i \sin^2 \varphi) \text{Var}(A)_i + (\langle A \rangle^2_i \sin^2 \varphi)}}.$$

(S23)

Therefore, the maximum of the SNR over $A_w$ finally turns out to be

$$\max \text{SNR}_p = g\eta(\varphi) \sqrt{N \left( \langle [\Omega, M] \rangle_D \right)^2 + \left| \langle [\Omega, M] \rangle_D \right|^2},$$

(S24)

where

$$\eta(\varphi) = \sqrt{\text{Var}(A)_i + \langle A \rangle^2_i \sin^2 \varphi}.$$

(S25)

An upper bound on $\text{SNR}_p$ can be obtained immediately from (S24) by the Robertson-Schrödinger uncertainty inequality [2]:

$$\max \text{SNR}_p \leq 2g\eta(\varphi) \sqrt{N \text{Var}(\Omega)_D}. $$

(S26)

III. COMPARISON OF SNR BETWEEN TWO TYPES OF WEAK MEASUREMENTS

A. Normal coherent state for the pointer

According to [S10], the SNR of a non-postselected weak measurement is

$$\text{SNR}_n = g \sqrt{N} \langle A \rangle_i \langle [\Omega, M] \rangle_D$$

(S27)

Since $\max \langle A \rangle_i = \lambda_{\text{max}}(A)$, the optimal SNR is

$$\max \text{SNR}_n = g \sqrt{N} \lambda_{\text{max}}(A) \langle [\Omega, M] \rangle_D,$$

(S28)

The ratio between the SNR of postselected and non-postselected weak measurements is

$$s = \frac{\max \text{SNR}_p}{\max \text{SNR}_n} = \frac{\sqrt{\text{Var}(A)_i \csc^2 \varphi + \langle A \rangle^2_i}}{|\lambda_{\text{max}}(A)|}.$$  

(S29)

Since $\text{Var}(A)_i = \langle A^2 \rangle_i - \langle A \rangle^2_i$ and $\csc^2 \varphi \geq 1$, we have $\text{Var}(A)_i \csc^2 \varphi + \langle A \rangle^2_i \geq \langle A^2 \rangle_i$, so when $|\Phi_i \rangle \rightarrow |\lambda_{\text{max}}(A)\rangle$, $s \geq 1$. The key question is when is $\csc^2 \varphi > 1$, so that $s > 1$?
Suppose $\Omega = q$, $M = p$. Then $\langle [\Omega, M] \rangle_D = i$, and

$$\csc^2 \varphi = 1 + (\langle q, p \rangle_D - 2 \langle q \rangle_D \langle p \rangle_D)^2.$$  

(S30)

In the Fock space, $q$ and $p$ can be represented by $q = \frac{1}{\sqrt{2}} (a + a^\dagger)$, $p = \frac{1}{\sqrt{2i}} (a - a^\dagger)$, so $\{q, p\} = i(a^\dagger a - a^2)$, and

$$\csc^2 \varphi = 1 - (\langle a^2 - a^2 \rangle_D + \langle a \rangle_D^2 - \langle a^\dagger a \rangle_D^2)^2.$$  

(S31)

When the initial pointer state $|D\rangle$ is a normal coherent state, $\langle a^\dagger a \rangle_D = \langle a \rangle_D^2$, $\langle a^2 \rangle_D = \langle a \rangle_D^2$, so $\csc^2 \varphi = 1$, which means that normal coherent states of the pointer cannot give postselected weak measurements any advantage in SNR compared with non-postselected weak measurements.

**B. Squeezed coherent state for the pointer**

Now turn to squeezed coherent states for the pointer. Suppose,

$$|D\rangle = |\xi, \alpha\rangle = \exp \frac{1}{2} (\xi^* a^2 - \xi a^\dagger^2) |\alpha\rangle,$$  

(S32)

where $\xi = re^{i\theta}$ is the squeeze parameter. Then it can be shown that

$$\langle a \rangle_D = \alpha \cosh r - \alpha^* e^{i\theta} \sinh r,$$

$$\langle a^\dagger \rangle_D = \langle a \rangle_D^*,$$

$$\langle a^2 \rangle_D = \alpha^2 \cosh^2 r + \alpha^* e^{2i\theta} \sinh^2 r - 2 |\alpha| \cosh r - e^{i\theta} \sinh r \cosh r,$$

$$\langle a^\dagger a \rangle_D = \langle a^2 \rangle_D^*. $$  

(S33)

So in this case,

$$\csc^2 \varphi = 1 + 4 (\sin \theta \sinh r \cosh r)^2.$$  

(S34)

When $r > 0$ and $\sin \theta \neq 0$, then $\csc^2 \varphi > 1$. Therefore the SNR of postselected weak measurements in this case can be enhanced beyond the SNR of non-postselected weak measurements. This implies that properly squeezed coherent states can increase the SNR of weak measurements by postselection while normal coherent states cannot.

**IV. QUANTUM FISHER INFORMATION OF TWO TYPES OF WEAK MEASUREMENTS**

In this section, we obtain the quantum Fisher information for postselected weak measurements and non-postselected weak measurements, respectively.

**A. The case of postselected weak measurements**

For a pure parameter-dependent state $|\psi_g\rangle$, the quantum Fisher information of estimating $g$ is

$$F^{(Q)} = 4 (\langle \partial_g \psi_g | \partial_g \psi_g \rangle - |\langle \psi_g | \partial_g \psi_g \rangle|^2).$$  

(S35)

We first consider postselected weak measurements. According to [S3] and [S6], the pointer state after the postselection on the system in a postselected weak measurement is

$$|D_f\rangle \approx e^{-i g A_w \Omega} |D\rangle \sqrt{1 + 2g \text{Im} A_w \langle D | \Omega | D \rangle},$$  

(S36)

so

$$|\partial_g D_f\rangle \approx -(i A_w \Omega + \langle \Omega | D \text{Im} A_w | D \rangle).$$  

(S37)
Therefore,

\begin{equation}
\langle \partial g D f | \partial g D f \rangle \approx |A_w|^2 \langle \Omega^2 \rangle_D - \text{Im}^2 A_w \langle \Omega \rangle_D^2.
\end{equation}

\( \langle D f | \partial g D f \rangle \approx -i \text{Re} A_w \langle \Omega \rangle_D. \) \hspace{1cm} (S38)

Hence, the quantum Fisher information of the post-interaction pointer state in this type of weak measurements is

\begin{equation}
F^{(Q)}_p \approx 4 P_s |A_w|^2 \text{Var}(\Omega)_D.
\end{equation}

\( \text{(S39)} \)

**B. The case of non-postselected weak measurements**

Next, we consider non-postselected weak measurements. In a non-postselected weak measurement, the post-interaction pointer state is usually a mixed state, since the pointer is entangled with the system by the interaction. The reduced density matrix of the pointer after the interaction is

\begin{equation}
\rho'_D = \text{tr}_S(\exp(-i g A \otimes \Omega) |\Phi_i \rangle \langle D| (\exp(g A \otimes \Omega)).
\end{equation}

(S40)

Suppose the eigenstates of \( A \) are \( |a_i\rangle, i = 1, \ldots, d \), where the \( a_i \)'s are the corresponding eigenvalues, and the initial state of the system is \( |\Phi_i \rangle = \sum_k \alpha_k |a_k\rangle \). Then \( \rho'_D \) becomes

\begin{equation}
\rho'_D = \sum_k |\alpha_k|^2 \exp(-i g a_k \Omega) |D\rangle \langle D| \exp(i g a_k \Omega).
\end{equation}

(S41)

Since \( g \ll 1 \), \( \exp(-i g a_k \Omega) \approx 1 - i g a_k \Omega \), and

\begin{align*}
\rho'_D & \approx \sum_k |\alpha_k|^2 (1 - i g a_k \Omega) |D\rangle \langle D| (1 + i g a_k \Omega) \\
& \approx \sum_k |\alpha_k|^2 (|D\rangle \langle D| - i g a_k \Omega, |D\rangle \langle D|)
\end{align*}

(S42)

where we have used \( \sum_k |\alpha_k|^2 = 1 \) and \( \sum_k |\alpha_k|^2 a_k = \langle A \rangle_i \).

Again, because \( g \ll 1 \),

\begin{equation}
|D\rangle \langle D| - i g \langle A \rangle_i |\Omega, |D\rangle \langle D| \approx \exp(-i g \langle A \rangle_i \Omega) |D\rangle \langle D| \exp(i g \langle A \rangle_i \Omega).
\end{equation}

(S43)

After the interaction, the pointer is approximately in a pure state:

\begin{equation}
|D_f \rangle \approx \exp(-i g \langle A \rangle_i \Omega) |D\rangle.
\end{equation}

(S44)

From Eq. \( \text{[S35]} \), we can immediately derive the quantum Fisher information of the post-interaction pointer state:

\begin{equation}
F^{(Q)}_n \approx 4 \langle A \rangle_i^2 \text{Var}(\Omega)_D.
\end{equation}

(S45)

[1] S. Pang, J. Dressel and T. A. Brun, Phys. Rev. Lett. 113, 030401 (2014).
[2] Robertson-Schrödinger uncertainty inequality: for any two Hermitian operators \( \Omega_1 \) and \( \Omega_2 \), there is

\begin{equation}
\Delta \Omega_1 \Delta \Omega_2 \geq \frac{1}{2} \sqrt{||[\Omega_1, \Omega_2]||^2 + ||[\Omega_1, \Omega_2]||^2}.
\end{equation}

(S46)

[3] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. 72, 3439 (1994).