Suppressing technical noise in weak measurements by entanglement

Shengshi Pang and Todd A. Brun

Phys. Rev. A **92**, 012120 — Published 27 July 2015

DOI: 10.1103/PhysRevA.92.012120

Suppressing technical noises in weak measurement by entanglement

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

Postselected weak measurement has aroused broad interest for its distinctive ability to amplify small physical quantities. However, the low postselection efficiency to obtain a large weak value has been a big obstacle to its application in practice, since it may waste resources, and reduce the measurement precision. An improved protocol was proposed in [Phys. Rev. Lett. **113**, 030401 (2014)] to make the postselected weak measurement dramatically more efficient by using entanglement. Such a protocol can increase the Fisher information of the measurement to approximately saturate the well-known Heisenberg limit. In this paper, we review the entanglement-assisted protocol of postselected weak measurement in detail, and study its robustness against technical noises. We focus on readout errors. Readout errors can greatly degrade the performance of postselected weak measurement, especially when the readout error probability is comparable to the postselection probability. We show that entanglement can significantly reduce the two main detrimental effects of readout errors: inaccuracy in the measurement result, and the loss of Fisher information. We extend the protocol by introducing a majority vote scheme to postselection to further compensate for readout errors. With a proper threshold, almost no Fisher information will be lost. These results demonstrate the effectiveness of entanglement in protecting postselected weak measurement against readout errors.

PACS numbers: 03.65.Ta, 03.65.Ud, 03.65.Ca, 03.67.Ac

I. INTRODUCTION

A quantum measurement is associated with an observable of the system in the standard von Neumann model. The effect of a quantum measurement is to stochastically project the system onto an eigenstate of the observable, and the reading from the measurement is the corresponding eigenvalue. Such a projective measurement usually requires the interaction between the system and the pointer to be strong, or the spread of the pointer wavefunction to be narrow, so that the final wavefunctions, translated by different eigenvalues of the observable, can be distinguished accurately.

In 1988, Aharonov, Albert, and Vaidman (AAV) [1] coined a quantum measurement protocol that violated the above conditions for projective measurements. In this protocol, the width of the pointer wave function is larger than the eigenvalue separation, or equivalently the interaction between the system and the pointer is weak, so that the final states of the pointer, kicked by different eigenvalues, have very large overlap with each other. AAV also introduced postselection of the system by measuring the system as well as the pointer, and retaining only those events where the system measurement produced a particular value. Such postselection destroys the correlation between the system and the pointer, and collapses the pointer to interfere the overlapping translations.

The AAV protocol of weak measurement with postselection seems trivial, but it can have a surprising physical effect: with a proper postselection of the system, the average translation of the pointer can be much larger than any eigenvalue of the observable, in sharp contrast to the standard projective quantum measurement. The mechanism behind this effect is that with the postselection of the system, the interference between the component wavefunctions, translated by different eigenvalues in the pointer state, may dramatically cancel the major part of the original wavefunction, resulting in a shift of the pointer that goes far beyond the eigenvalue spectrum of the observable.

The large shift of the pointer can be approximated as a linear amplification of the otherwise weak interaction strength. Such a linear amplification can be characterized by a quantity called the *weak value* [1] (which will be introduced in detail in Sec. II). This value can have formally divergent behavior when the postselected state of the system asymptotically becomes orthogonal to the initial state of the system. (Of course, in practice, the amplification effect cannot be infinitely large. This has been extensively investigated in recent years [2–5].)

There was some controversy over the condition for the validity of the weak value after the birth of AAV weak measurement, but it was soon clarified [6]. In addition, unlike the eigenvalues of an observable, the weak value is generally a complex value, with its real and imaginary parts playing different roles in the amplification effect [7].

The measurement of weak values has been realized experimentally [8], and the amplification effect has been found useful in observing weak physical effects in many real experiments, including the spin Hall effect of light [9–11], optical beam deflection [12–18], optical frequency shift [19], optical phase shift [20, 21], temperature shift [22], temporal shift [23–25], etc. More experimental protocols have also been proposed [23, 26–36]. Moreover, weak measurements have been realized on systems other

* shengshp@usc.edu
† tbrun@usc.edu

than optical systems, including superconducting circuits [37–39], NMR [40], among others. For a comprehensive review of postselected weak measurement and weak value amplification, we refer the readers to [41–44].

When postselected weak measurement is used to amplify small physical quantities, a large weak value is usually desired. However, this will lead to a low postselection efficiency, since the weak value is approximately reciprocal to the square root of the postselection probability. Low postselection efficiency may result in a waste of resources and reduce the Fisher information of the measurement if the failed postselections are discarded, and consequently cancel the advantage of weak value amplification. In fact, this has led to a recent hot debate: whether postselecting the system and discarding the unselected events can ever increase the precision of the measurement.

Some studies have suggested that the weak value amplification can produce higher precision in weak measurements [14, 27, 31, 34], while others found the opposite results [45–47]. More extensive work on this issue has made clear that the Fisher information of weak measurements cannot be increased by postselection if the unselected events are discarded [46, 48, 49], because the discarded events take away some Fisher information [50], and the distribution of postselection probabilities can also carry information [48, 49, 51]. Interestingly, however, if all these sources of Fisher information are taken into account, the total Fisher information can saturate the Heisenberg limit in some cases, even with seemingly classical resources [51, 52]. Moreover, weak value amplification can give an advantage in suppressing technical noise [53] (although not all types of noises can be suppressed [54]), or even use technical noises to enhance the sensitivity [55]. A brief review of the recent controversy over weak value amplification can be found in [56].

Because of the above problems of low postselection efficiency, an important goal in the practice of weak value amplification is to raise the postselection probability. Some efforts have been made on this. For example, it was proposed that by recycling unpostselected photons, almost every photon can eventually be successfully postselected [33].

In a recent study [57], it was noticed that for a given weak value, the choice of pre- and postselection of the system to realize it is usually not unique, so there is some freedom to maximize the postselection probability. Alternatively, if the postselection probability is fixed, there is some freedom to maximize the magnitude of the weak value. Maximizing the postselection probability or weak value can dramatically improve the resource usage or performance of the weak value amplification. And the result of either optimization shows that the loss of Fisher information in the discarded events can be made as small as the order of the interaction strength, which is usually negligible in the weak value approximation. This implies that postselected weak measurement can offer technical advantages [53] at almost no cost in Fisher information.

Based on these optimization results for the postselection efficiency and weak value, an improved protocol of postselected weak measurement assisted by entanglement was proposed in [57]. The protocol uses entangled systems rather than uncorrelated systems. If there are n systems entangled initially, the postselection probability scales as n^2, while if the n systems are uncorrelated, the postselection probability can only scale linearly in n. Thus, entanglement between the systems can bring a marked increase (of order n) in the postselection probability. An important consequence of this increase in postselection probability is that the Fisher information can also be correspondingly raised by the order of n (because the Fisher information is proportional to the size of the data sample), and approach the Heisenberg limit, which is the upper bound on estimation precision achievable in quantum metrology [58].

This paper builds on [57] to detail the entanglement-assisted protocol of weak measurement, and its advantages in improving the metrological performance of weak measurement. Furthermore, since technical noise is inevitable in real experimental devices, we will study the influence of technical noise on this protocol, and show the robustness of this protocol against the noise. The main technical noise we will consider is readout error in postselecting the system.

Readout errors mix successful postselection events with unsuccessful ones. Since the shift of the pointer resulting from unsuccessful postselections is much smaller than from successful postselections, and the sensitivity of the pointer states in the former case is also much lower, mixing them will bring errors to the measurement result and reduce the precision of the measurement. Moreover, as will be shown, readout errors may be more detrimental in postselected weak measurements than in other quantum measurements. When the postselection probability is very small, even a low rate of readout errors may cause severe problems. So it is necessary to suppress the influence of readout errors to make postselected weak measurements more reliable in practice.

In this paper, we will analyze the robustness of postselected weak measurement against readout errors, and show that the use of entanglement can correct the deviation of measurement result caused by readout errors with an extremely high success rate. Furthermore, it will be shown that entanglement can recover part of the Fisher information reduced by readout errors. Introducing an appropriate measurement threshold strategy can decrease the Fisher information loss to nearly zero. These results suggest that entanglement, combined with a measurement threshold strategy, can effectively suppress the effect of readout errors on postselected weak measurements.

This paper is organized as follows. First, in Sec. II, we briefly introduce the weak value theory of postselected weak measurement and how it leads to the amplification effect. Then, in Sec. III, we study the optimization of postselection to maximize the postselection probability given the weak value, or to maximize the weak value given the postselection probability. Sec. IV is devoted to introducing entanglement-assisted weak measurement based on the optimization result. Sec. V gives a detailed study of the Fisher information of postselected weak measurement, and shows that the Fisher information can saturate the Heisenberg limit with the assistance of entanglement. A qubit example is given in Sec. VI to illustrate the entanglement-assisted protocol, and verify the saturation of Heisenberg limit. Finally, in Sec. VII, we investigate the influence of readout errors in detail, and introduce a measurement threshold scheme to protect the Fisher information against readout errors.

II. REVIEW OF WEAK VALUE FORMALISM

In a standard quantum measurement, the measurement results are eigenvalues of a system observable, and the system collapses to the eigenstate of the observable corresponding to the measurement result. A typical model to realize this standard quantum measurement is

$$H_{\text{int}} = g\hat{A}\otimes\hat{F}\delta(t-t_0), \tag{1}$$

where $\hat{A}$ and $\hat{F}$ are observables of the system and the pointer respectively, and g characterizes the strength of the interaction. Suppose the initial state of the system is $|\Psi_i\rangle$, and and initial state of the pointer is $|D\rangle$, then the system and the pointer are coupled by the interaction, and evolve to an entangled state

$$|\Phi\rangle = \exp(-\mathrm{i}g\hat{A}\otimes\hat{F})|\Psi_i\rangle|D\rangle. \tag{2}$$

If we expand $|\Psi_i\rangle$ along the eigenstates of $\hat{A}$, $|\Phi\rangle$ can be written as

$$|\Phi\rangle = \sum_k c_k|a_k\rangle\exp(-\mathrm{i}ga_k\hat{F})|D\rangle, \tag{3}$$

where a_k, $|a_k\rangle$ are eigenvalues and eigenstates of $\hat{A}$, and c_k are the expansion coefficients of $|\Psi_i\rangle$ in the basis of $\{|a_k\rangle\}$.

Different $\exp(-\mathrm{i}ga_k\hat{F})$ in (3) transform $|D\rangle$ into different states. If $|D\rangle$ is properly chosen so that the overlaps between $\exp(-\mathrm{i}ga_k\hat{F})|D\rangle$ are sufficiently small, the $\exp(-\mathrm{i}ga_k\hat{F})|D\rangle$ can be distinguished with a low error probability then, and the measurement on the pointer will make the system collapse to a state that is close to an eigenstate of $\hat{A}$. For example, suppose $\hat{F}$ is the momentum operator $\hat{p}$, then $(-\mathrm{i}ga_k\hat{p})$ is a translation operator in the position space of the pointer, and if one can measure the position of the pointer, the measurement results will be ga_k, and the system will collapse to $|a_k\rangle$. If g is set to 1, it will lead to the formalism of standard projective quantum measurement.

A major innovation by AAV in weak measurements was introducing postselection of the system. This seemingly minor change turns out to give some surprising results that are dramatically different from ordinary quantum measurements.

In detail: if the system is postselected to the state $|\Psi_f\rangle$ after the interaction, the pointer then collapses to the (unnormalized) state

$$|D_f\rangle = \langle\Psi_f|\exp(-\mathrm{i}g\hat{A}\otimes\hat{F})|\Psi_i\rangle|D\rangle. \tag{4}$$

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

$$\begin{aligned}|D_f\rangle &\approx \langle\Psi_f|(1-\mathrm{i}g\hat{A}\otimes\hat{F})|\Psi_i\rangle|D\rangle\\ &= \langle\Psi_f|\Psi_i\rangle(1-\mathrm{i}gA_w\hat{F})|D\rangle,\end{aligned} \tag{5}$$

where A_w is the weak value,

$$A_w = \frac{\langle\Psi_f|\hat{A}|\Psi_i\rangle}{\langle\Psi_f|\Psi_i\rangle}. \tag{6}$$

If g is so small that $gA_w \ll 1$, $|D_f\rangle$ can be rewritten as

$$|D_f\rangle \approx \exp(-\mathrm{i}gA_w\hat{F})|D\rangle. \tag{7}$$

So one can see that in the presence of postselection of the system, the pointer is shifted by roughly gA_w (in the representation of a complementary observable of $\hat{F}$). In sharp contrast to ordinary measurement, the shift of the pointer in this case can be much larger than any ga_k, because A_w can be much larger than 1 when $|\langle\Psi_f|\Psi_i\rangle| \ll 1$.

Note that A_w can be complex, and in this case $\exp(-\mathrm{i}gA_w\hat{F})$ is not just a simple translation operator. In fact, it can be decomposed to the product of a translation operator (corresponding to the real part of A_w) and a state reduction operator (corresponding to the imaginary part of A_w). Jozsa gave a very detailed study of complex weak value in [7], and analyzed the role of the real and imaginary parts of the weak value. He showed that if the pointer observable $\hat{F}$ is the momentum $\hat{p}$, then the shifts in the average position and momentum of the pointer are, respectively,

$$\begin{aligned}\langle\Delta\hat{q}\rangle &= g\mathrm{Re}A_w + g\mathrm{Im}A_w(m\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Var}_{\hat{q}}),\\ \langle\Delta\hat{p}\rangle &= 2g\mathrm{Im}A_w\mathrm{Var}_{\hat{p}},\end{aligned} \tag{8}$$

where $\hat{q}$ and $\hat{p}$ are the position and momentum operators of the pointer.

That result can be generalized to a more general form. Suppose one measures an observable $\hat{M}$ on the pointer after postselecting the system. The average shift of the pointer is

$$\langle\Delta\hat{M}\rangle_f = \frac{\langle D_f|\hat{M}|D_f\rangle}{\langle D_f|D_f\rangle} - \langle\hat{M}\rangle_{|D\rangle}. \tag{9}$$

From Eq. (5), one can get

$$\begin{aligned}\langle D_f|\hat{M}|D_f\rangle &\approx |\langle\Psi_f|\Psi_i\rangle|^2(\langle\hat{M}\rangle_{|D\rangle} + \mathrm{i}g\mathrm{Re}A_w\langle[\hat{F},\hat{M}]\rangle_{|D\rangle}\\ &\quad + g\mathrm{Im}A_w\langle\{\hat{F},\hat{M}\}\rangle_{|D\rangle}),\\ \langle D_f|D_f\rangle &\approx |\langle\Psi_f|\Psi_i\rangle|^2(1+2g\mathrm{Im}A_w\langle\hat{F}\rangle_{|D\rangle}),\end{aligned} \tag{10}$$

so plugging (10) into (9) produces

$$\langle \Delta \hat{M} \rangle_f \approx g \mathrm{Im} A_w (\langle \{\hat{F}, \hat{M}\} \rangle_{|D\rangle} - 2\langle \hat{F} \rangle_{|D\rangle} \langle \hat{M} \rangle_{|D\rangle}) + \mathrm{i} g \mathrm{Re} A_w \langle [\hat{F}, \hat{M}] \rangle_{|D\rangle}. \tag{11}$$

Note that if one plugs $\hat{F} = \hat{p}$ and $\hat{M} = \hat{q}$, $\hat{p}$ into (11), the result (8) can be immediately recovered.

Eqs. (8) and (11) imply that the shift of the pointer is roughly proportional to the weak value A_w when $g \ll 1$. Since A_w can be much larger than 1 when $\langle \Psi_f | \Psi_i \rangle \ll 1$, the shift of the pointer can be treated as an amplification of g by the weak value A_w. This is the origin of the amplification effect in postselected weak measurements. This amplification effect has been widely used in experiments to measure small parameters, as reviewed in the introduction.

III. OPTIMIZATION OF WEAK MEASUREMENTS

A shortcoming of postselected weak measurements that can be seen immediately from (6) is that a large weak value A_w requires a very small overlap between the initial state and the postselected state of the system. This implies that the postselection probability must be very low, because when $g \ll 1$, the success probability of a postselection is approximately

$$P_s \approx |\langle \Psi_f | \Psi_i \rangle|^2. \tag{12}$$

Therefore, an important problem in a practical application of weak measurement amplification is to improve the postselection probability as much as possible while the weak value is still kept large.

From the definition of the weak value (6), when a weak value A_w is fixed, the possible choices of the system initial state $|\Psi_i\rangle$ and postselected state $|\Psi_f\rangle$ to realize this weak value is not unique, and different $|\Psi_i\rangle$ and $|\Psi_f\rangle$ may give different postselection probabilities. This provides the possibility to optimize the choice of pre- and postselections of the system to maximize the the postselection efficiency for a given weak value.

Alternatively, if the postselection probability (12) is given, different pre- and postselections of the system can also produce different weak values, which will lead to different amplification abilities for the parameter g. So there exist optimal pre- and postselections of the system to produce the maximum weak value for a given postselection probability.

The significance of optimizing weak measurements to maximize either the postselection probability or the weak value is obvious: one can reduce the resources needed to give a desired amplification effect, or one can make the best use of the given resources to produce the maximum amplification effect. So it is useful for practical applications of weak value amplification to optimize the performance.

In this section, we study these two ways of optimizing weak measurements in detail. Besides the significance mentioned above, the optimizations derived in this section will also be the foundation of the entanglement-assisted weak measurements that we will introduce in the next section.

A. Maximum postselection probability given the weak value

In this subsection, we study the first optimization problem introduced above: that is, to maximize the postselection probability over all possible pre- and postselections of the system for a given weak value.

To solve this problem directly using the constraint on the fixed weak value A_w in maximizing the postselection probability is rather difficult; so in order to utilize this condition, we first convert it to another more accessible form. Note that Eq. (6) can be rewritten as

$$\langle \Psi_f | (\hat{A} - A_w) | \Psi_i \rangle = 0, \tag{13}$$

so the constraint of fixed A_w can be reinterpreted as $|\Psi_f\rangle$ being orthogonal to $(\hat{A} - A_w)|\Psi_i\rangle$. With this new form of the condition, the optimization of $|\langle \Psi_f | \Psi_i \rangle|^2$ can be much simplified.

From a geometrical point of view, it is not difficult to verify that $|\Psi_f\rangle$ should be parallel to the component of $|\Psi_i\rangle$ in the subspace orthogonal to $(\hat{A} - A_w)|\Psi_i\rangle$ (which we denote as $\mathcal{V}^\perp$ below) when $|\langle \Psi_f | \Psi_i \rangle|^2$ is maximized. Therefore, we can first decompose the initial state of the system along $(\hat{A} - A_w)|\Psi_i\rangle$ and its orthogonal subspace $\mathcal{V}^\perp$,

$$|\Psi_i\rangle = \frac{(\hat{A} - A_w)|\Psi_i\rangle \langle \Psi_i | (\hat{A} - A_w^*) | \Psi_i \rangle}{\langle \Psi_i | (\hat{A} - A_w^*)(\hat{A} - A_w) | \Psi_i \rangle} + \left(|\Psi_i\rangle - \frac{(\hat{A} - A_w)|\Psi_i\rangle \langle \Psi_i | (\hat{A} - A_w^*) | \Psi_i \rangle}{\langle \Psi_i | (\hat{A} - A_w^*)(\hat{A} - A_w) | \Psi_i \rangle} \right), \tag{14}$$

and then the optimal $|\Psi_f\rangle$ can be obtained:

$$|\Psi_f\rangle \propto |\Psi_i\rangle - \frac{(\hat{A} - A_w)|\Psi_i\rangle \langle \Psi_i | (\hat{A} - A_w^*) | \Psi_i \rangle}{\langle \Psi_i | (\hat{A} - A_w^*)(\hat{A} - A_w) | \Psi_i \rangle}. \tag{15}$$

Hence, the maximum postselection probability for the given weak value A_w is

$$\max P_s = \left\| |\Psi_i\rangle - \frac{(\hat{A} - A_w)|\Psi_i\rangle \langle \Psi_i | (\hat{A} - A_w^*) | \Psi_i \rangle}{\langle \Psi_i | (\hat{A} - A_w^*)(\hat{A} - A_w) | \Psi_i \rangle} \right\|^2 = \frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{\langle \Psi_i | \hat{A}^2 | \Psi_i \rangle - 2\langle \Psi_i | \hat{A} | \Psi_i \rangle \mathrm{Re} A_w + |A_w|^2}, \tag{16}$$

where $\mathrm{Var}(\hat{A})_{|\Psi_i\rangle} = \langle \Psi_i | \hat{A}^2 | \Psi_i \rangle - \langle \Psi_i | \hat{A} | \Psi_i \rangle^2$ is the variance of $\hat{A}$ in the state $|\Psi_i\rangle$.

For the purposes of weak value amplification, we usually desire $|A_w|$ to be larger than any eigenvalue of $\hat{A}$, $|A_w| \gg \max|\lambda(\hat{A})|$, which implies that

$$\begin{aligned} |A_w| &\gg \langle\Psi_i|\hat{A}|\Psi_i\rangle, \\ |A_w| &\gg \sqrt{\langle\Psi_i|\hat{A}^2|\Psi_i\rangle}. \end{aligned} \tag{17}$$

Therefore, the maximum P_s can be approximated as

$$\max P_s \approx \frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{|A_w|^2}, \tag{18}$$

for a large A_w.

B. Maximum weak value given the postselection probability

In this subsection, we solve the second optimization problem that was introduced at the beginning of this section: that is, to maximize the weak value over all possible pre- and postselections of the system for a given postselection probability. Of course, the weak value is generally complex, so we will focus on maximizing the amplitude of the weak value.

Suppose the postselection probability is P_s. Since the phase of the post-selected state $|\Psi_f\rangle$ does not affect the weak value A_w, $|\Psi_f\rangle$ can be written as

$$|\Psi_f\rangle = \sqrt{P_s}|\Psi_i\rangle + \sqrt{1-P_s}|\Psi_i^\perp\rangle, \tag{19}$$

where $|\Psi_i^\perp\rangle$ is a state orthogonal to $|\Psi_i\rangle$. This implies that we can write the weak value in Eq. (6) as

$$A_w = \langle\Psi_i|\hat{A}|\Psi_i\rangle + \sqrt{\frac{1-P_s}{P_s}}\langle\Psi_i^\perp|\hat{A}|\Psi_i\rangle. \tag{20}$$

Now to maximize A_w is just to maximize $\langle\Psi_i^\perp|\hat{A}|\Psi_i\rangle$ over $|\Psi_i^\perp\rangle$.

Similar to the maximization procedure in the last subsection, the maximum $\langle\Psi_i^\perp|\hat{A}|\Psi_i\rangle$ can be achieved when $|\Psi_i^\perp\rangle$ is parallel to the component of $\hat{A}|\Psi_i\rangle$ in the complementary subspace orthogonal to $|\Psi_i\rangle$, so

$$\begin{aligned} |\langle\Psi_i^\perp|\hat{A}|\Psi_i\rangle|_{\max} &= \left\| \hat{A}|\Psi_i\rangle - |\Psi_i\rangle\langle\Psi_i|\hat{A}|\Psi_i\rangle \right\| \\ &= \sqrt{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}. \end{aligned} \tag{21}$$

Therefore, the largest weak value that can be obtained from the initial state $|\Psi_i\rangle$ with a given post-selection probability P_s is

$$\max|A_w| = \langle\Psi_i|\hat{A}|\Psi_i\rangle + \sqrt{\frac{1-P_s}{P_s}\,\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}. \tag{22}$$

For a large $|A_w|$, $P_s \ll 1$, so the first term in (22) can be neglected, thus,

$$\max|A_w| \approx \sqrt{\frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{P_s}}. \tag{23}$$

The results of both optimization problems, Eqs. (18) and (23), indicate that the maximum postselection probability or the maximum weak value are proportional to the variance (or the square root of the variance) of the observable $\hat{A}$ in the initial state of the system. This observation leads directly to the entanglement-assisted weak measurement protocol that we introduce in the next section.

IV. ENTANGLEMENT-ASSISTED WEAK MEASUREMENT

As mentioned at the end of the last section, both the maximum postselection probability (18) and the maximum weak value (23) are proportional to the variance of $\hat{A}$ (or its square root) in the initial state of the system. Since the variance of an observable scales differently with the number of the systems in an entangled state than in an uncorrelated state, this observation leads to a new weak measurement protocol that we introduce in this section. We shall see how entanglement can assist the weak measurement either to save resources or to improve the amplification.

A. Uncorrelated systems

As a reference example, we first consider the case where the systems have no correlation between each other initially. Suppose we have n systems. If these n systems are initially in a product state $|\Psi_i^1\rangle \otimes \cdots \otimes |\Psi_i^n\rangle$, and the postselections are $|\Psi_f^1\rangle, \cdots, |\Psi_f^n\rangle$, which give the same weak value to each individual system, i.e.,

$$\frac{\langle\Psi_f^k|\hat{A}|\Psi_i^k\rangle}{\langle\Psi_f^k|\Psi_i^k\rangle} = A_w,\ k = 1, \cdots, n, \tag{24}$$

then, when $|\langle\Psi_f^k|\Psi_i^k\rangle| \ll 1$, the probability to have at least one successful event in postselecting these n systems is

$$\begin{aligned} P_s^{(n)} &= 1 - \prod_{k=1}^{n}\left(1 - |\langle\Psi_f^k|\Psi_i^k\rangle|^2\right) \\ &\approx \sum_{k=1}^{n} |\langle\Psi_f^k|\Psi_i^k\rangle|^2. \end{aligned} \tag{25}$$

Now, if the choice of pre- and postselections for each system maximizes $|\langle\Psi_f^k|\Psi_i^k\rangle|^2$, i.e., $|\langle\Psi_f^k|\Psi_i^k\rangle|^2 = P_s^{(1)}$, where $P_s^{(1)}$ is the maximal postselection probability for a single system, then

$$P_s^{(n)} \approx nP_s^{(1)}. \tag{26}$$

This implies that the postselection efficiency increases linearly with n when the systems are initially uncorrelated.

In fact, the linear scaling of $P_s^{(n)}$ with n in (26) is the best scaling that can be obtained with initially uncorrelated systems. When $|\Psi_i\rangle$ is a product state of n systems, say $|\Psi_i^{(n)}\rangle = |\Psi_i^1\rangle \otimes \cdots \otimes |\Psi_i^n\rangle$, it can be verified that

$$\mathrm{Var}(\hat{A}^{(n)})_{|\Psi_i^{(n)}\rangle} = \mathrm{Var}(\hat{A}_1)_{|\Psi_i^1\rangle} + \cdots + \mathrm{Var}(\hat{A}_n)_{|\Psi_i^n\rangle}. \quad (27)$$

When each $|\Psi_i^k\rangle$ maximizes $\mathrm{Var}(\hat{A}_k)$, then

$$\mathrm{Var}(\hat{A}^{(n)})_{|\Psi_i^{(n)}\rangle} = n\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}, \quad (28)$$

where we omitted the subscript in $\hat{A}$ and the superscript in $|\psi\rangle$ on the right side of (28) since each individual system has the same $\hat{A}$ and $|\psi\rangle$. According to Eq. (18), the maximum postselection probability is proportional to the variance of $\hat{A}$ in the initial state of the system, so (28) implies that if the initial state of the n systems is a product state, the postselection probability $P_s^{(n)}$ at most can scale linearly with n.

B. Entangled systems

Now let us remove the constraint that the systems are uncorrelated, and see whether the postselection probability can be improved. For n systems, we first need to generalize the observable $\hat{A}$. The observable $\hat{A}$ becomes the sum of n single-system observables in this case,

$$\hat{A}^{(n)} = \hat{A}_1 + \cdots \hat{A}_n, \quad (29)$$

where we use the superscript (n) to denote the n-system observable explicitly.

Now suppose we have n systems, and they are initially prepared in the following entangled state:

$$|\Psi_i^{(n)}\rangle = \alpha|a_{\max}\rangle^{\otimes n} + \beta|a_{\min}\rangle^{\otimes n}. \quad (30)$$

Then it can be obtained directly that

$$\begin{aligned}\mathrm{Var}(\hat{A}^{(n)})_{|\Psi_i^{(n)}\rangle} = {} & n^2[|\alpha|^2 a_{\max}^2 + |\beta|^2 a_{\min}^2 \\ & - (|\alpha|^2 a_{\max} + |\beta|^2 a_{\min})^2].\end{aligned} \quad (31)$$

One can immediately see that with the entangled state (30) (and $\alpha \neq 0$, $\beta \neq 0$), the scaling of $\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}$ becomes quadratic with n, which is higher than with a product $|\Psi_i\rangle$ by order n. So the maximum postselection probability can be increased by order n in this case.

Is quadratic scaling with n optimal when entanglement is used? And what entangled state of the system maximizes the factor before n^2 in the variance $\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}$? To answer these two questions, let us recall that for an arbitrary Hermitian operator $\hat{\Xi}$, its maximum variance over all possible states $|\psi\rangle$ is

$$\max_{|\psi\rangle} \mathrm{Var}(\hat{\Xi})_{|\psi\rangle} = \frac{1}{4}(\xi_{\max} - \xi_{\min})^2, \quad (32)$$

where $\xi_{\max}$ and $\xi_{\min}$ are the maximum and minimum eigenvalues of $\hat{\Xi}$ respectively, and the maximum variance is attained when

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\xi_{\max}\rangle + \mathrm{e}^{\mathrm{i}\theta}|\xi_{\min}\rangle), \quad (33)$$

where $|\xi_{\max}\rangle$ and $|\xi_{\min}\rangle$ are the corresponding eigenstates of $\hat{\Xi}$, and $\mathrm{e}^{\mathrm{i}\theta}$ is an arbitrary phase.

Applying this fact to the observable $\hat{A}^{(n)}$ in (29), we obtain

$$\max \mathrm{Var}(\hat{A}^{(n)})_{|\Psi_i\rangle} = n^2 \max \mathrm{Var}(\hat{A}), \quad (34)$$

and

$$\max \mathrm{Var}(\hat{A}) = \frac{1}{4}(a_{\max} - a_{\min})^2. \quad (35)$$

This follows because

$$\lambda_{\max}(\hat{A}^{(n)}) = na_{\max},\ \lambda_{\min}(\hat{A}^{(n)}) = na_{\min}, \quad (36)$$

where

$$\lambda_{\max}(\hat{A}) = a_{\max},\ \lambda_{\min}(\hat{A}) = a_{\min}. \quad (37)$$

From Eq. (18), one sees that the maximum postselection probability over all pre- and postselections of the n systems is

$$P_s^{(n)} \approx n^2 P_s^{(1)}, \quad (38)$$

which is increased by order n compared to the uncorrelated state (26).

What price do we pay for such an improvement of postselection probability? Note that the eigenstates of $\hat{A}^{(n)}$ with eigenvalues (36) are

$$|\lambda_{\max}\rangle = |a_{\max}\rangle^{\otimes n},\ |\lambda_{\min}\rangle = |a_{\min}\rangle^{\otimes n}. \quad (39)$$

So according to (33), the initial state of the n systems should be

$$|\Psi_i^{(n)}\rangle = \frac{1}{\sqrt{2}}(|a_{\max}\rangle^{\otimes n} + \mathrm{e}^{\mathrm{i}\theta}|a_{\min}\rangle^{\otimes n}), \quad (40)$$

which is an entangled state. Therefore, to improve the postselection efficiency we need entanglement in the initial state of the n systems.

To obtain the maximum postselection probability, the postselected state of the n systems should also be carefully chosen. According to (15) and (40), in order to maximize the postselection probability, the system should be postselected to the following state after the weak interaction:

$$\begin{aligned}|\Psi_f^{(n)}\rangle \propto {} & -(na_{\min} - A_w^*)|a_{\max}\rangle^{\otimes n} \\ & + \mathrm{e}^{\mathrm{i}\theta}(na_{\max} - A_w^*)|a_{\min}\rangle^{\otimes n}.\end{aligned} \quad (41)$$

When $|A_w| \gg \max\{n|a_{\max}|, n|a_{\min}|\}$, $|\Psi_f^{(n)}\rangle$ can be simplified to

$$|\Psi_f^{(n)}\rangle \propto \mathrm{e}^{\frac{n\Delta}{A_w^*}}|a_{\max}\rangle^{\otimes n} - \mathrm{e}^{\mathrm{i}\theta}|a_{\min}\rangle^{\otimes n}, \quad (42)$$

where $\Delta = a_{\max} - a_{\min}$.

Similarly, we can also use the maximally entangled state (40) as the initial state of the n systems to increase the weak value with the postselection probability fixed. The only difference from the previous protocol is the choice of the postselected state of the systems. When the postselection probability is fixed to P_s, the postselected state contains two components: one is $\sqrt{P_s}|\Psi_i^{(n)}\rangle$, and the other is $\sqrt{1-P_s}|\Psi_i^{(n)\perp}\rangle$. According to Sec. III B, for the optimal postselection, $|\Psi_i^{(n)\perp}\rangle$ should be proportional to the component of $\hat{A}^{(n)}|\Psi_i^{(n)}\rangle$ in the subspace orthogonal to the state $|\Psi_i^{(n)}\rangle$, i.e.,

$$|\Psi_i^{(n)\perp}\rangle \propto \hat{A}^{(n)}|\Psi_i^{(n)}\rangle - |\Psi_i^{(n)}\rangle\langle\Psi_i^{(n)}|\hat{A}^{(n)}|\Psi_i^{(n)}\rangle, \quad (43)$$

which turns out to be

$$|\Psi_i^{(n)\perp}\rangle = \frac{1}{\sqrt{2}}(|a_{\max}\rangle^{\otimes n} - e^{i\theta}|a_{\min}\rangle^{\otimes n}). \quad (44)$$

Therefore, the optimal postselected state is

$$\begin{aligned}|\Psi_f^{(n)}\rangle = &\left(\sqrt{\frac{P_s}{2}} + \sqrt{\frac{1-P_s}{2}}\right)|a_{\max}\rangle^{\otimes n} \\ &+ e^{i\theta}\left(\sqrt{\frac{P_s}{2}} - \sqrt{\frac{1-P_s}{2}}\right)|a_{\min}\rangle^{\otimes n}.\end{aligned} \quad (45)$$

The postselected state $|\Psi_f^{(n)}\rangle$ (either (41) or (45)) is also an entangled state. Generally speaking, postselecting n systems in an entangled state is very nontrivial. In Sec. VI, we show how to achieve this kind of postselection by simple quantum circuits for qubits. That method can be generalized to higher dimensional systems.

V. FISHER INFORMATION OF WEAK MEASUREMENT

Precision is one of the most important benchmarks for the performance of a measurement. Since weak measurement can amplify small parameters by postselecting the system in addition to measuring the pointer, it has long been speculated that the postselected weak measurements can increase the precision of measuring small parameters. However, since failed postselection events comprise a large fraction of total events, the precision can also be significantly reduced by low postselection efficiency. It is possible that the increase of the precision by weak value amplification may be canceled by low efficiency. Therefore, whether postselection can really increase the precision of weak measurement is controversial, and it is important to make clear how well weak measurement with postselection can do in the metrology of small parameter estimation. This has become a hot topic of recent research.

In this section, we study this problem in detail. We compute the Fisher information, a widely used metric for the precision of parameter estimation, for general weak measurements with postselection of the system, and use the results from the last section to obtain the maximum Fisher information for a given weak value or a fixed postselection probability. With the assistance of entanglement among the systems, the Fisher information may be increased approximately to the Heisenberg limit, which is the upper bound for quantum Fisher information, and the loss of Fisher information in the failed postselection events can be reduced to the order of the small parameter, which is negligible in the regime of the weak value approximation. So the performance of entanglement-assisted weak value amplification essentially achieves the optimum for quantum metrology.

A. Background of Fisher information

We first introduce the Fisher information. The precision of estimating a parameter is usually quantified by the variance of the estimate. But it is often not easy to compute the exact variance of an estimate, since the estimate itself often does not have an analytical solution for many estimation strategies. Fortunately, the Cramér-Rao relation [59] gives a lower bound for the variance of an unbiased estimate, quantified by the Fisher information.

For a probability distribution $p_g(x)$ dependent on a parameter g, the minimum statistical error of estimating g from $p_g(x)$ satisfies

$$\langle\delta g^2\rangle \geq \frac{1}{nI_g} + \langle\delta g\rangle^2, \quad (46)$$

where n is the number of sample data, and I_g is the Fisher information defined as

$$I_g = \int_X \frac{(\partial_g p_g(x))^2}{p_g(x)} dx, \quad (47)$$

where X is the region that x belongs to. If the estimation strategy is unbiased, the estimate bias $\langle\delta g\rangle$ will vanish, and the statistical error of the estimate is lower bounded by the reciprocal of the Fisher information. It can be proved that the lower bound (46) can be saturated in the limit $n \to \infty$ when the estimation uses the maximum likelihood strategy.

For a quantum state $|\Phi_g\rangle$ dependent on a parameter g, the method to estimate g is to measure many copies of $|\Phi_g\rangle$, and estimate g from the distribution of measurement results. Suppose the measurement is described by a POVM $\{\hat{E}_1, \cdots, \hat{E}_r\}$, where

$$\hat{E}_i \geq 0, \text{ and } \sum_i \hat{E}_i = I. \quad (48)$$

Then the probability of obtaining the i-th result is

$$p(i) = \langle\Phi_g|\hat{E}_i|\Phi_g\rangle. \quad (49)$$

Obviously, the probability of measurement results $p(i)$ is dependent on g, so from the distribution $p(i)$ one can estimate the parameter g. And the Fisher information of estimation can be obtained by plugging (49) into (47).

Since there are many different choices of measurement on $|\Phi_g\rangle$, which lead to different Fisher informations of estimating g, there exists a maximum of the Fisher information over all choices of measurement. This maximum Fisher information is called the *quantum Fisher information* [60, 61], and is found to be

$$I_g^{(Q)} = 4\big(\langle\partial_g\Phi_g|\partial_g\Phi_g\rangle - |\langle\Phi_g|\partial_g\Phi_g\rangle|^2\big). \tag{50}$$

In (50), the dependence on the choice of measurement has vanished, and the quantum Fisher information $I_g^{(Q)}$ is determined solely by the state $|\Phi_g\rangle$.

A common task in quantum metrology is to estimate some multiplicative parameter g of a Hamiltonian in the form gH. In this case, one usually prepares a quantum system in some initial state $|\Phi\rangle$ and let it evolve under the Hamiltonian gH for some time t. The final state of the system is $\exp(-itg\hat{H})|\Phi\rangle$. Then one can do a measurement on the system, and when the measurement is optimized, the quantum Fisher information is [58]

$$I_g^{(Q)} = 4\mathrm{Var}(\hat{H})_{|\Phi\rangle}, \tag{51}$$

which is entirely determined by the variance of the Hamiltonian $\hat{H}$ in the initial state $|\Phi\rangle$.

B. General result for the Fisher information of weak measurements

With the above background knowledge of quantum Fisher information, we go on to study the precision of postselected weak measurements in this subsection. Our central focus is to investigate whether the competition between the amplification by the weak value and the reduction by the low postselection probability leads to a gain or a loss of the Fisher information, and how much the gain or the loss is. To achieve this aim, we compare the quantum Fisher information of estimating g with and without postselection of the system.

In a weak measurement with (1) as the interaction Hamiltonian and $|\Psi_i\rangle$, $|D\rangle$ as the respective initial states of the system and pointer, the whole system evolves to the joint state $\exp(-\mathrm{i}g\hat{A}\otimes\hat{F})|\Psi_i\rangle\otimes|D\rangle$ after the weak interaction. If there is no postselection of the system after the interaction, then according to Eq. (51), the quantum Fisher information is

$$I_g^{(Q)} = 4\Big[\langle\hat{A}^2\rangle_{|\Psi_i\rangle}\langle\hat{F}^2\rangle_{|D\rangle} - \big(\langle\hat{A}\rangle_{|\Psi_i\rangle}\langle\hat{F}\rangle_{|D\rangle}\big)^2\Big]. \tag{52}$$

Now suppose we perform a projective measurement on the system in order to make a postselection. This measurement will produce d independent outcomes corresponding to some orthonormal basis $\{|\Psi_f^k\rangle\}_{k=1}^d$, where d is the dimension of the system. In the linear response regime with $g \ll 1$, each of these outcomes corresponds to a postselection of the system, and collapses the pointer to the state

$$|D_{fk}\rangle \approx \big(\hat{I} - \mathrm{i}gA_w^{(k)}\hat{F}\big)|D\rangle, \tag{53}$$

with success probability $P_s^{(k)} \approx |\langle\Psi_f^{(k)}|\Psi_i\rangle|^2$, and

$$A_w^{(k)} = \frac{\langle\Psi_f^k|\hat{A}|\Psi_i\rangle}{\langle\Psi_f^k|\Psi_i\rangle}. \tag{54}$$

Then according to (50), the Fisher information in each of the collapsed states $|D_{fk}\rangle$ after postselecting the system to $|\Psi_f^k\rangle$ is

$$\begin{aligned}I_g^{(k)} \approx\ & 4\,P_s^{(k)}|A_w^{(k)}|^2\Big[\mathrm{Var}(\hat{F})_{|D\rangle} \\ & - \langle\hat{F}^2\rangle_{|D\rangle}\big(2g\mathrm{Im}A_w^{(k)}\langle\hat{F}\rangle_{|D\rangle} + g^2|A_w^{(k)}|^2\langle\hat{F}^2\rangle_{|D\rangle}\big)\Big].\end{aligned} \tag{55}$$

It is important to observe that if we add the information from all d postselections, we obtain

$$\sum_{k=1}^{d} I_g^{(k)} \approx 4\langle\hat{A}^2\rangle_{|\Psi_i\rangle}\mathrm{Var}(\hat{F})_{|D\rangle} - O(g), \tag{56}$$

where we have used

$$\sum_k P_s^{(k)}|A_w^{(k)}|^2 = \langle\hat{A}^2\rangle_{|\Psi_i\rangle}. \tag{57}$$

With the condition $\langle\hat{F}\rangle_{|D\rangle} = 0$, $\mathrm{Var}(\hat{F})_{|D\rangle} = \langle\hat{F}^2\rangle_{|D\rangle}$, then (56) saturates the maximum in (52) up to a small corrections of order g, which indicates that the measurement on the system does not lose information by itself, but rather redistributes and concentrates the information about g in the final pointer states $\{|D_{fk}\rangle\}_{k=1}^d$.

In a postselected weak measurement, if we postselect the system in the state $|\Psi_f^k\rangle$, we discard all events in which the system does not collapse to $|\Psi_f^k\rangle$. So for such a postselected weak measurement, the Fisher information where the pointer is initially in an unbiased state (i.e., $\langle\hat{F}\rangle_{|D\rangle} = 0$) is

$$I_g^{(k)} \approx 4P_s^{(k)}|A_w^{(k)}|^2\langle\hat{F}^2\rangle_{|D\rangle}\big(1 - g^2|A_w^{(k)}|^2\langle\hat{F}^2\rangle_{|D\rangle}\big). \tag{58}$$

It is worth mentioning that Ref. [50] derived results about the total Fisher information of postselected weak measurement similar to Eqs. (55) and (58), though in a slightly different notation, and Ref. [48] obtained a result similar to (56), including the classical Fisher information of the postselected result distribution.

C. Maximum Fisher information of postselected weak measurements

Having obtained the general result for quantum Fisher information of a postselected weak measurement in Eq.

(58), we now go on to consider maximizing the quantum Fisher information from a postselected weak measurement, and investigate whether, and how much, Fisher information is lost by discarding the failed postselection events. This is currently the subject of hot debate.

From Eq. (58) we see that the dependence of the Fisher information $I_g^{(k)}$ on the initial and postselected states of the system is determined by the postselection probability $P_s^{(k)}$ and the weak value $A_w^{(k)}$. To maximize the Fisher information $I_g^{(k)}$, we use the results on the maximum postselection probability given the weak value, or maximum weak value given the postselection probability, that were obtained in Sec. III.

From Sec. III, we know that when the weak value is sufficiently large, i.e., $|A_w| \gg \max\lambda(\hat{A})$, then the maximum P_s and the maximum $|A_w|$ can be approximated as

$$\begin{aligned} \max P_s &\approx \frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{|A_w|^2}, \qquad \text{with } A_w \text{ fixed}, \\ \max|A_w| &\approx \sqrt{\frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{P_s}}, \qquad \text{with } P_s \text{ fixed}. \end{aligned} \tag{59}$$

Now, we can plug either equation of (59) into (58), and obtain

$$\max I_g^{(k)} \approx 4\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}\langle\hat{F}^2\rangle_{|D\rangle}\big(1 - g^2|A_w^{(k)}|^2\langle\hat{F}^2\rangle_{|D\rangle}\big). \tag{60}$$

Comparing this maximum Fisher information for weak measurement with postselection to that without postselection in (52), it follows that

$$\max I_g^{(k)} \approx I_g^{(Q)}\frac{\mathrm{Var}(\hat{A})_{|\Psi_i\rangle}}{\langle\hat{A}^2\rangle_{|\Psi_i\rangle}}\big(1 - |gA_w^{(k)}|^2\langle\hat{F}^2\rangle_{|D\rangle}\big), \tag{61}$$

where $I_g^{(Q)}$ is the global quantum Fisher information with an unbiased pointer ($\langle\hat{F}\rangle_{|D\rangle} = 0$). And it is almost equal to $I_g^{(Q)}$ (52) if the system is initially unbiased ($\langle\hat{A}\rangle_{|\Psi_i\rangle} = 0$).

This implies that postselection redistributes the quantum Fisher information between the system and the pointer, and with the optimal choice of pre- and postselection of the system, it can concentrate nearly all the Fisher information into a single (but very improbable) pointer state. The remaining very small amount of information is distributed among the failed postselection events, and could be retrieved in principle by measuring the respective collapsed pointer states. The pointer state corresponding to successful postselection of the system suffers an overall reduction factor of $\mathrm{Var}(\hat{A})/\langle\hat{A}^2\rangle$ which is 1 for unbiased system states, as well as a tiny loss $|gA_w^{(1)}|^2\langle\hat{F}^2\rangle_{|D\rangle}$. However, most weak value amplification experiments operate in the linear response regime $g|A_w^{(1)}|\langle\hat{F}^2\rangle_{|D\rangle}^{\frac{1}{2}} \ll 1$, so this remaining loss is negligible. Moreover, the overall reduction can be further compensated by extracting information from the postselection probability distribution [51].

This is quite a surprising result because it implies that one can approximately saturate the global optimal bound of Fisher information (52) by measuring only the very rare postselected pointer state while the remaining much more probable outcomes are discarded, although the optimal bound (52) cannot be exactly reached [48, 50, 51]. The unlikely postselections can also offer an advantage in practice: in measuring small signals in the face of experimental imperfections, it can be easier to see rare large events than frequent small ones. This property of postselected weak measurement makes it a broadly useful technique for estimating small parameters within the linear response regime [53].

D. Saturation of Heisenberg limit

A well-known upper bound on the quantum Fisher information is the Heisenberg limit, which sets the ultimate upper bound for the sensitivity of quantum parameter estimation, and demonstrates that the parameter estimation precision by quantum measurements can scale as n^{-1} (or equivalently, quantum Fisher information can scale quadratically with n), if n quantum systems are coupled. This is higher than the standard quantum limit (SQL) (or the classical limit) $n^{-\frac{1}{2}}$, by order $\sqrt{n}$.

An interesting question in weak value amplification is whether weak measurement with postselection can achieve Heisenberg-limited scaling in precision. As we showed in the last subsection, the quantum Fisher information of postselected weak measurement can be approximately promoted to the global maximum quantum Fisher information $I_g^{(Q)}$ by optimized pre- and postselections of the system. The global quantum Fisher information can generally achieve the Heisenberg limit using entangled systems (and pointers), so we would expect that the Fisher information of postselected weak measurement can also reach the Heisenberg limit.

We can straightforwardly verify this idea by exploiting the results that were obtained in the previous sections. Similar to standard quantum metrology, when the n systems are uncorrelated the Fisher information of a postselected weak measurement scales like the standard quantum limit, i.e., n, while with entanglement among the systems, the Fisher information can be boosted to scale like the Heisenberg limit n^2.

In Eq. (60), it was shown that the maximum Fisher information of a postselected weak measurement is proportional to the variance of the system observable $\hat{A}$ in the initial state of the system. Therefore, the scaling of the variance determines the scaling of the Fisher information.

Suppose we have n systems. From Sec. IV A and IV B, we know that the variance of the total observable $\hat{A}^{(n)}$ scales linearly with n when the n systems are initially uncorrelated, and scales quadratically with n when the n systems are initially entangled. Therefore, we can immediately conclude that the maximum Fisher information of a postselected weak measurement with n systems can indeed scale like the Heisenberg limit n^2, and the necessary ingredient to reach this limit is entanglement between the systems initially.

The achievability of the Heisenberg limit by using entanglement can also be understood in another way. It is known that the total Fisher information of estimating a parameter from a sample of data is proportional to the size of the sample. Since failed postselection events are discarded in postselected weak measurements, the total Fisher information is proportional to the postselection probability. We know from Sec. III that the postselection probability scales linearly with n when the n systems are initially uncorrelated, and scales quadratically with n when the n systems are initially entangled. Therefore, the total Fisher information can scale as n^2 if the initial state of the n systems is entangled.

The above simple result again verifies the previous conclusion that with optimized pre- and postselections of the system, the Fisher information of a postselected weak measurement can approximately reach the global optimal bound, and the loss of Fisher information by discarding the failed postselection events can be negligible.

VI. QUBIT EXAMPLE

To illustrate the new protocol of entanglement-assisted weak measurement, we give an example with qubits in this section. For clarity, we will focus on the protocol for increasing the postselection probability with a fixed weak value from now on. The case of increasing the weak value with a fixed postselection probability can be derived straightforwardly.

A. Protocol

Suppose we use n qubits as systems and let them couple to a common pointer qubit. The interaction Hamiltonian between each system qubit and the pointer qubit is

$$H_{\mathrm{int}} = \varphi\hat{\sigma}_z \otimes \hat{\sigma}_x,\ \varphi \ll 1. \tag{62}$$

In the entanglement-assisted weak measurement scheme, the initialization step prepares the n qubits in a maximally entangled state:

$$|\Psi_i^{(n)}\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}). \tag{63}$$

This can be achieved by inputting $n-1$ qubits in the state $|0\rangle$ and one qubit in the state $|1\rangle$ into the circuit in Fig. VI A. A sequence of $n-1$ CNOT gates entangles the n qubits, and produces the desired maximally entangled state (63).

Next, the n qubits are subject to the weak interaction with the common pointer qubit (62). To simulate this

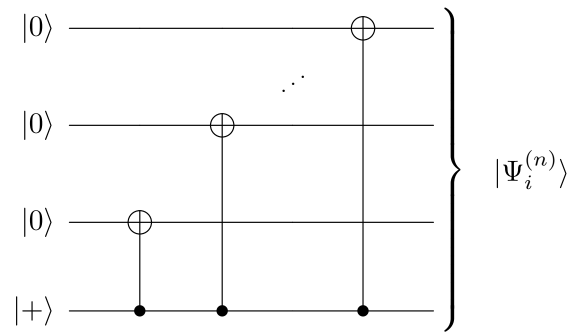


Figure 1. Quantum circuit to initialize the n system qubits. The qubits are prepared in the maximally entangled state $|\Psi_i\rangle = (|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$ by a sequence of CNOT gates.

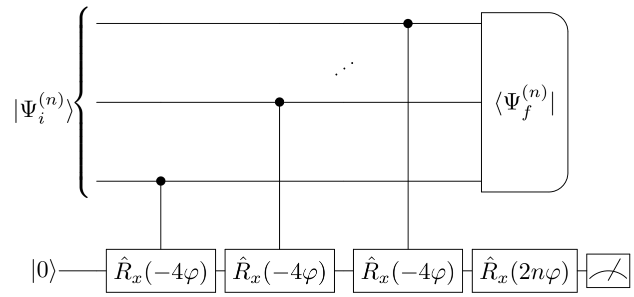


Figure 2. Quantum circuit equivalent to the weak interaction $H_{\mathrm{int}} = \varphi\hat{\sigma}_z \otimes \hat{\sigma}_x$, $\varphi \ll 1$, between n system qubits and one common pointer qubit.

interaction, note that the unitary interaction under H_{int} can be written as

$$\begin{aligned}\hat{U} &= \exp(-\mathrm{i}\varphi\hat{\sigma}_z \otimes \hat{\sigma}_x)\\ &= |0\rangle\langle 0| \otimes \exp(-\mathrm{i}\varphi\hat{\sigma}_x) + |1\rangle\langle 1| \otimes \exp(\mathrm{i}\varphi\hat{\sigma}_x)\\ &= (|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \exp(2\mathrm{i}\varphi\hat{\sigma}_x))(I \otimes \exp(-\mathrm{i}\varphi\hat{\sigma}_x)),\end{aligned} \tag{64}$$

so $\hat{U}$ can be realized with a control-$\hat{R}_x(-4\varphi)$ rotation followed by a $\hat{R}_x(2\varphi)$ gate on the pointer qubit, where $\hat{R}_x(\theta)$ [62] is defined as

$$\hat{R}_x(\theta) = \exp\Big(-\mathrm{i}\frac{\theta}{2}\hat{\sigma}_x\Big). \tag{65}$$

Since the control-$\hat{R}_x(-4\varphi)$ rotation and the $\hat{R}_x(2\varphi)$ rotation commute, all the $\hat{R}_x(2\varphi)$ gates can be delayed until after the last control-$\hat{R}_x(-4\varphi)$ rotation. So the total interaction $U^{\otimes n}$ is equivalent to by n control-$\hat{R}_x(-4\varphi)$ rotations between the system qubits and the pointer qubit followed by an $\hat{R}_x(2n\varphi)$ gate on the pointer.

Therefore, the weak interaction in the weak measurement using n system qubits and one common pointer qubit can be implemented by the circuit in Fig. VI A.

After the weak interactions, the n system qubits are postselected to the state $|\Psi_f^{(n)}\rangle$ (41), which in this case turns out to be

$$|\Psi_f^{(n)}\rangle \propto (n + A_w^*)|0\rangle^{\otimes n} + (n - A_w^*)|1\rangle^{\otimes n}. \tag{66}$$

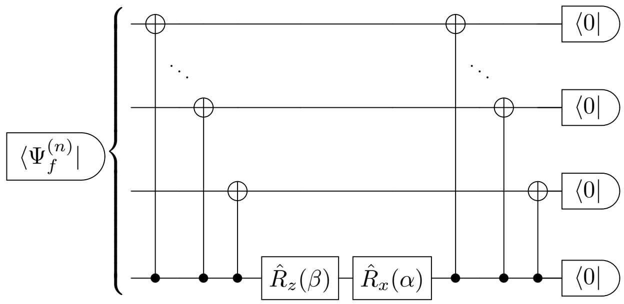


Figure 3. Quantum circuit to postselect the n system qubits in the entangled state $|\Psi_f\rangle$ (66), which is in essence the reverse process of preparing $|\Psi_f\rangle$ from $|0\rangle^{\otimes n}$.

Such a postselection realizes a given weak value A_w.

Postselection of n system qubits in an entangled state $|\Psi_f^{(n)}\rangle$ (66) is usually not easy, but it can be realized as the inverse procedure for a preparation. Note that $|\Psi_f^{(n)}\rangle$ can be written as

$$|\Psi_f^{(n)}\rangle = \hat{V}|0\rangle^{\otimes n}, \tag{67}$$

where $\hat{V}$ is a unitary transformation that turns $|0\rangle^{\otimes n}$ to $|\Psi_f^{(n)}\rangle$. $\hat{V}$ can be realized by CNOT gates and single qubit gates that are similar to the preparation of $|\Psi_i^{(n)}\rangle$.

There are many different choices of $\hat{V}$ to realize (67). One convenient choice among them is to transform the subspace spanned by $\{|0\rangle^{\otimes n}, |1\rangle^{\otimes n}\}$ to itself, i.e., the subspace spanned by $\{|0\rangle^{\otimes n}, |1\rangle^{\otimes n}\}$ is invariant under $\hat{V}$. The advantage of such a choice is that there are only two possible postselected states of the n system qubits, $|0\rangle^{\otimes n}$ and $|1\rangle^{\otimes n}$, and these two states have the largest Hamming distance, which will be helpful to fighting against readout errors that will be discussed in Sec. VII.

Such a $\hat{V}$ can be written as

$$\hat{V}: \begin{cases} |0\rangle^{\otimes n} \longrightarrow |\Psi_f^{(n)}\rangle, \\ |1\rangle^{\otimes n} \longrightarrow |\Psi_f^{(n)\perp}\rangle, \end{cases} \tag{68}$$

where

$$|\Psi_f^{(n)\perp}\rangle \propto (n + A_w)|1\rangle^{\otimes n} - (n - A_w)|0\rangle^{\otimes n}. \tag{69}$$

Then, the postselection of the n system qubits to $|\Psi_f^{(n)}\rangle$ can be decomposed into the reverse unitary transformation $\hat{V}^\dagger$ followed by a postselection in the state $|0\rangle^{\otimes n}$. It can be implemented by the circuit in Fig. VI A, where

$$\begin{aligned} \alpha &= -2\arctan\sqrt{\frac{n^2 + |A_w|^2 - 2n\mathrm{Re}A_w}{n^2 + |A_w|^2 + 2n\mathrm{Re}A_w}}, \\ \beta &= -\frac{\pi}{2} - \arg\frac{n - A_w^*}{n + A_w^*}. \end{aligned} \tag{70}$$

In the following, we compute the composite state of the n system qubits and the pointer qubit before the postselection, and the probability of successful postselection. These results will be useful in the remainder of the paper.

Suppose the initial state of the n system qubits is $|\Psi_i\rangle = (|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$. Then after the weak interaction, the joint state of the system qubits and the pointer qubit is

$$\begin{aligned} |\Phi\rangle &= \hat{V}^\dagger \exp(-\mathrm{i}\varphi(\hat{\sigma}_{1z} + \cdots + \hat{\sigma}_{nz}) \otimes \hat{\sigma}_x)|\Psi_i\rangle|D\rangle \\ &\propto (|0\rangle^{\otimes n}\langle\Psi_f^{(n)}| + |1\rangle^{\otimes n}\langle\Psi_f^{(n)\perp}|) \\ &\quad (|0\rangle^{\otimes n}\mathrm{e}^{-\mathrm{i}n\varphi\hat{\sigma}_x} + |1\rangle^{\otimes n}\mathrm{e}^{\mathrm{i}n\varphi\hat{\sigma}_x})|D\rangle. \end{aligned} \tag{71}$$

According to (66) and (69), $|\Phi\rangle$ can be simplified to

$$\begin{aligned} |\Phi\rangle \propto\ & |0\rangle^{\otimes n}((n + A_w)\mathrm{e}^{-\mathrm{i}n\varphi\hat{\sigma}_x} + (n - A_w)\mathrm{e}^{\mathrm{i}n\varphi\hat{\sigma}_x})|D\rangle \\ &+ |1\rangle^{\otimes n}((-n + A_w^*)\mathrm{e}^{-\mathrm{i}n\varphi\hat{\sigma}_x} + (n + A_w^*)\mathrm{e}^{\mathrm{i}n\varphi\hat{\sigma}_x})|D\rangle. \end{aligned} \tag{72}$$

So, when the n system qubits are postselected in $|0\rangle^{\otimes n}$, the pointer state collapses to

$$|D_{f,0}\rangle = (n\cos n\varphi - \mathrm{i}A_w \sin n\varphi\hat{\sigma}_x)|D\rangle, \tag{73}$$

and when the n system qubits are postselected in $|1\rangle^{\otimes n}$, the pointer qubit collapses to

$$|D_{f,1}\rangle = (A_w^* \cos n\varphi + \mathrm{i}n \sin n\varphi\hat{\sigma}_x)|D\rangle. \tag{74}$$

The postselection probabilities for $|0\rangle^{\otimes n}$ and $|1\rangle^{\otimes n}$ can be worked out from (72) respectively:

$$\begin{aligned} p_0^{(n)} &= \frac{\eta_0}{\eta_0 + \eta_1}, \\ p_1^{(n)} &= \frac{\eta_1}{\eta_0 + \eta_1}, \end{aligned} \tag{75}$$

where the superscripts (n) denote there are n entangled qubits, and

$$\begin{aligned} \eta_0^{(n)} &= n^2\cos^2 n\varphi + |A_w|^2\sin^2 n\varphi + n\mathrm{Im}A_w \sin 2n\varphi\langle\hat{\sigma}_x\rangle_D, \\ \eta_1^{(n)} &= |A_w|^2\cos^2 n\varphi + n^2\sin^2 n\varphi - n\mathrm{Im}A_w \sin 2n\varphi\langle\hat{\sigma}_x\rangle_D. \end{aligned} \tag{76}$$

B. Fisher information and Heisenberg limit

In this subsection, we calculate the Fisher information for the qubit example. We will see that with entanglement between the system qubits, the Fisher information of the pointer qubit can indeed reach the Heisenberg limit and saturate the Cramér-Rao bound.

Based on the the final pointer state and the postselection probability obtained in the last subsection, we can compute the quantum Fisher information of an entanglement-assisted postselected weak measurement. To simplify the computation, we assume that $n|\varphi| \ll 1$, $|A_w\varphi| \ll 1$, and $n^2|\varphi| \ll |A_w|$ which usually hold in the weak value approximation. Then $|D_{f0}\rangle$ and $|D_{f1}\rangle$ can be simplified to

$$\begin{aligned} |D_{f0}\rangle &\approx (I - \mathrm{i}A_w\varphi\hat{\sigma}_x)|D\rangle, \\ |D_{f1}\rangle &\approx \Big(I + \mathrm{i}n^2\frac{\varphi}{A_w^*}\hat{\sigma}_x\Big)|D\rangle, \end{aligned} \tag{77}$$

and

$$p_0^{(n)} \approx \frac{n^2}{|A_w|^2 + n^2}, \, p_1^{(n)} \approx \frac{|A_w|^2}{|A_w|^2 + n^2}. \tag{78}$$

According to (77), $|\partial_\varphi D_{f0}\rangle \approx -\mathrm{i}A_w \hat{\sigma}_x |D\rangle$, so from the definition of quantum Fisher information (50), and taking the postselection probability into consideration, the quantum Fisher information of the final pointer state when the n qubits are all postselected to $|0\rangle$ becomes

$$I_g^{(|0\rangle^{\otimes n})} = \frac{4n^2|A_w|^2(1-|A_w|^2\varphi^2)}{n^2 + |A_w|^2}. \tag{79}$$

When $|A_w|\varphi \ll 1$ and $|A_w| \gg n$,

$$I_g^{(|0\rangle^{\otimes n})} \approx 4n^2. \tag{80}$$

This shows that the Heisenberg limit is approximately attained in this case, and Fisher information is only lost to order φ^2. Since $\varphi \ll 1$ in the weak coupling regime, the loss of Fisher information is negligible. Therefore, the Fisher information of a postselected weak measurement can indeed approach the Heisenberg limit with optimal pre- and postselections of the systems, and the final state of the pointer qubit possesses almost all the Fisher information of the phase φ. This verifies the result in Sec. V D.

VII. READOUT ERRORS

In the previous sections, we introduced entanglement-assisted weak measurement and studied its performance in metrology. We now turn to issues that will arise in practical implementations.

Errors are inevitable in any practical application of a quantum protocol. They can be caused by noise in the environment, or by technical imperfections. Numerous ways have been invented to fight against errors, and systematic theories, such as quantum error correction code and dynamical decoupling [63], have been developed to utilize them.

In this section, we study a typical kind of error in postselected weak measurements: readout errors. Readout error can result from noise in the environment and technical imperfections in the experimental devices. We focus on the readout errors that occur in the postselection stage. This kind of error can significantly influence the weak measurement protocol, by distorting the postselected results. As we shall see later, even when the probability of a readout error is very low, the disturbance to the postselection measurement can be dramatic. So correcting this type of error is necessary in postselected weak measurements.

Readout errors influence postselection results in two main ways. First, they may read some failed postselections as successful ones, which can bring errors to the statistics of the postselection results. Second, they may read some successful postselections as failed ones, which will reduce the postselection efficiency.

Below, we start from the qubit example in Sec. VI and analyze the effects of readout errors in postselection. We will show that an initially entangled state of the n system qubits can dramatically increase the robustness of weak measurement, which demonstrates another advantage of entanglement. We will also study the loss rate of successful postselections, and its complementarity relation with the error rate. In addition, the effect of readout errors on the Fisher information will be considered. To fight against this type of error, we introduce a majority vote scheme to reduce both the error rate and the loss rate in the postselection results. This simple trick can eliminate almost all loss of Fisher information caused by readout errors.

A. Error in postselection results

In this subsection, we analyze the first effect of readout errors on a postselected weak measurement: that is, the relative error rate of the successful postselection results. (We will omit the adjective "relative" when there is no ambiguity in the context.) The second effect of readout errors, i.e., the loss of correct postselection results, will be discussed in the next subsection.

Suppose the probability of mistaking $|0\rangle$ for $|1\rangle$ is $q_{0\to1}$ and the probability of mistaking $|1\rangle$ for $|0\rangle$ is $q_{1\to0}$. Then in an ordinary weak measurement, the probability of obtaining a $|0\rangle$ from reading a system qubit is

$$p(|0\rangle) = p_0^{(1)}(1-q_{0\to1}) + p_1^{(1)}q_{1\to0}. \tag{81}$$

The component $p_1^{(1)}q_{1\to0}$ is the readout error which identifies $|1\rangle$ as $|0\rangle$, so the relative error rate in the postselection results is

$$p(\text{error}||0\rangle) = \frac{p_1^{(1)}q_{1\to0}}{p_0^{(1)}(1-q_{0\to1}) + p_1^{(1)}q_{1\to0}}. \tag{82}$$

Both $q_{1\to0}$ and $q_{0\to1}$ are usually small. However, when $A_w \gg 1$, $p(\text{error}||0\rangle)$ can become very large in some cases. To see this, use $p_1^{(1)} = 1 - p_0^{(1)}$ in (82), then (82) becomes

$$p(\text{error}||0\rangle) = \frac{(1-p_0^{(1)})q_{1\to0}}{p_0^{(1)}(1-q_{0\to1}) + (1-p_0^{(1)})q_{1\to0}}. \tag{83}$$

The error rate $p(\text{error}||0\rangle)$ can range from 0 to 1 when p_0 goes from 1 to 0.

In particular, if $p_0^{(1)} = q_{0\to1} = q_{1\to0}$, then

$$p(\text{error}||0\rangle) = \frac{1}{2}; \tag{84}$$

and if $p_0^{(1)} \ll q_{1\to0}$, then

$$p(\text{error}||0\rangle) \to 1. \tag{85}$$

These imply that the postselected weak measurement is very sensitive to readout errors, especially when the readout error probability is comparable to the postselection probability.

In a postselected weak measurement, the postselection probability is usually very small, so even a low probability of readout errors may lead to a significant error in the postselection results. Therefore, correcting readout errors or suppressing their influence is crucial to practical applications of postselected weak measurement.

Now let's consider the entanglement-assisted protocol of weak measurement, and see whether the relative error rate in the postselection results can be improved.

In this protocol, the probability of correctly identifying a $|0\rangle^{\otimes n}$ in the postselection is $(1-q_{0\to1})^n$, and the probability of mistaking a $|1\rangle^{\otimes n}$ for $|0\rangle^{\otimes n}$ is $q_{1\to0}^n$, so the total probability of reading out $|0\rangle^{\otimes n}$ from the n system qubits after the weak coupling is

$$p(|0\rangle^{\otimes n}) = p_0^{(n)}(1-q_{0\to1})^n + p_1^{(n)} q_{1\to0}^n. \tag{86}$$

Similar to the above, the erroneous proportion of the postselection results is $p_1 q_{1\to0}^n$, so the relative error probability in the postselection results $|0\rangle^{\otimes n}$ is

$$p(\mathrm{error}||0\rangle^{\otimes n}) = \frac{p_1^{(n)} q_{1\to0}^n}{p_0^{(n)}(1-q_{0\to1})^n + p_1^{(n)} q_{1\to0}^n}. \tag{87}$$

In Fig. 4, we plot the relative error probability $p(\mathrm{error}||0\rangle^{\otimes n})$ versus the weak value A_w for different numbers n of entangled systems. We assume $q_{1\to0}$ and $q_{0\to1}$ to be the same for simplicity. As n increases, the relative error $p(\mathrm{error}||0\rangle^{\otimes n})$ dramatically decreases.

Note that $p(\mathrm{error}||0\rangle^{\otimes n})$ can be rewritten as

$$p(\mathrm{error}||0\rangle^{\otimes n}) = \left(1 + \frac{p_0^{(n)}}{p_1^{(n)}}\left(\frac{1-q_{0\to1}}{q_{1\to0}}\right)^n\right)^{-1}. \tag{88}$$

From (75), (76), when $n\varphi A_w \ll 1$, we have

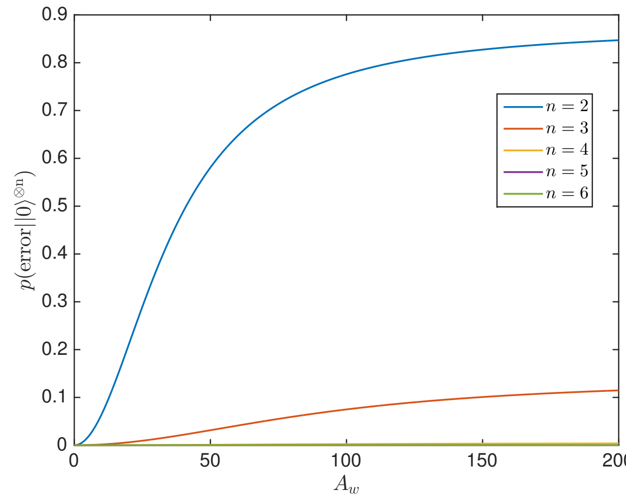
$$\frac{p_0^{(n)}}{p_1^{(n)}} \approx \frac{n^2}{|A_w|^2}. \tag{89}$$

Since $q_{0\to1}, q_{1\to0}$ are usually very small, if n further satisfies that $\frac{1-q_{0\to1}}{q_{1\to0}} \gg \sqrt[n]{\frac{|A_w|^2}{n^2}}$, then Eq. (88) can be approximately simplified to

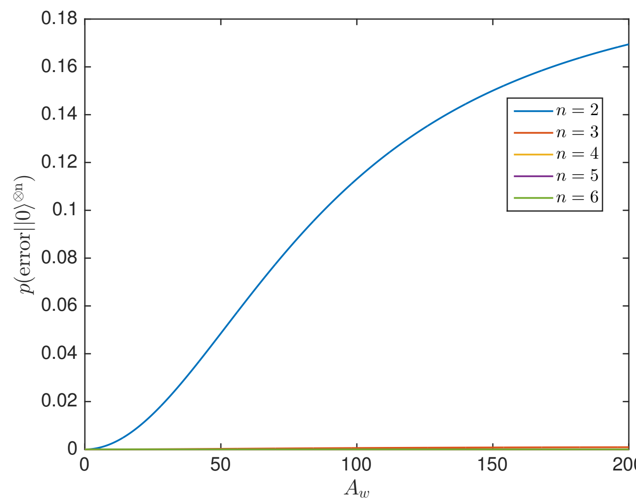
$$p(\mathrm{error}||0\rangle^{\otimes n}) \approx |A_w|^2 n^{-2}\left(\frac{1-q_{0\to1}}{q_{1\to0}}\right)^{-n}. \tag{90}$$

This means that entanglement between the initial system qubits can reduce the relative error rate caused by readout errors super-exponentially with the number of entangled qubits! It implies how efficiently entanglement can improve the robustness of postselected weak measurement against the readout errors.

To see this super-exponential decay of the relative error rate more clearly, the relative error rate $p(\mathrm{error}||0\rangle^{\otimes n})$ against n is plotted in Fig. 5 for different weak values. In the figure, when n is small, $p(\mathrm{error}||0\rangle^{\otimes n})$ does not drop rapidly. This is because $\left(\frac{1-q_{0\to1}}{q_{1\to0}}\right)^n$ is not large enough and the constant term 1 in (88) cannot be neglected. Nevertheless, when n grows large, $p(\mathrm{error}||0\rangle^{\otimes n})$ decreases much faster in the plot, which is what (90) predicted.



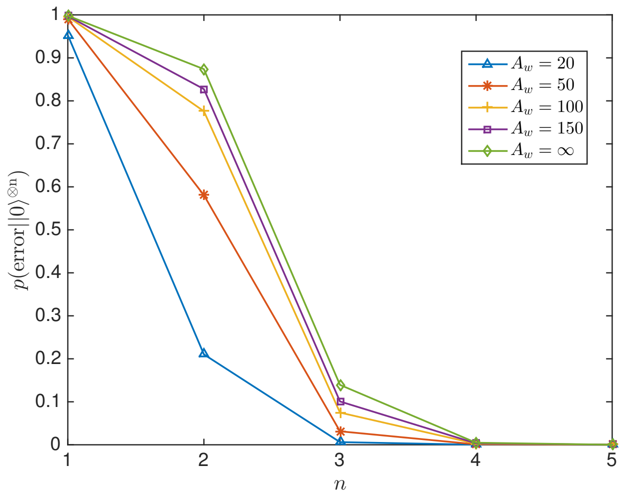
(a) $q_{1\to0} = q_{0\to1} = 0.05$



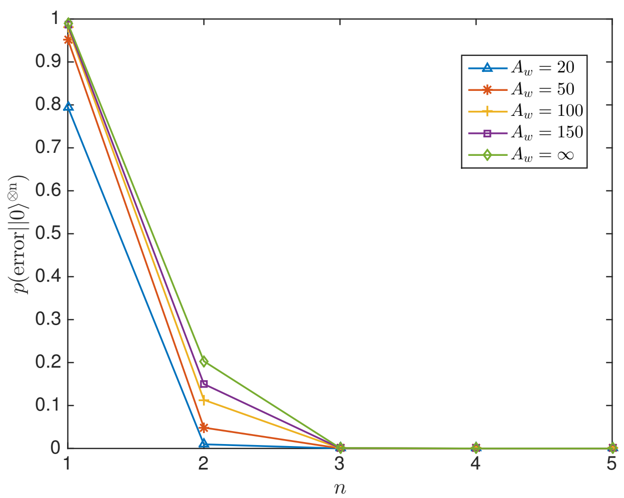
(b) $q_{1\to0} = q_{0\to1} = 0.01$

Figure 4. (Color online) Plot of the relative error probability $p(\mathrm{error}||0\rangle^{\otimes n})$ versus the weak value A_w for different n. The lines from left to right in each figure are for $n = 2, 3, 4, 5, 6$, respectively. (4a) When $q_{1\to0} = q_{0\to1} = 0.05$, the $p(\mathrm{error}||0\rangle^{\otimes n})$ for $n = 4, 5, 6$ are much lower than $n = 2, 3$, and they almost overlap since they are very close to each other. (4b) Similarly, when $q_{1\to0} = q_{0\to1} = 0.01$, the $p(\mathrm{error}||0\rangle^{\otimes n})$ for $n = 3, 4, 5, 6$ are much lower than $n = 2$, and they almost overlap with each other.

An interesting phenomenon can be observed in Fig. 4: when A_w grows large, the relative error rate $p(\mathrm{error}||0\rangle^{\otimes n})$ does not approach 1, and the line of $p(\mathrm{error}||0\rangle^{\otimes n})$ approaches a plateau as $A_w \to \infty$. This



(a) $q_{1\to 0} = q_{0\to 1} = 0.05$



(b) $q_{1\to 0} = q_{0\to 1} = 0.01$

Figure 5. (Color online) This figure plots the relative error probability $p(\text{error}||0\rangle^{\otimes n})$ versus the number of entangled qubits n for different weak values A_w. It shows that the relative error rate $p(\text{error}||0\rangle^{\otimes n})$ can decrease very fast with n, which implies the advantage of entanglement in suppressing the relative error rate.

can be explained from the above results. When $A_w \to \infty$, the leading terms of η_0 and η_1 in Eq. (76) are $|A_w|^2 \sin^2 n\varphi$ and $|A_w|^2 \cos^2 n\varphi$ respectively, so

$$\begin{aligned}\lim_{A_w\to\infty} p(\text{error}||0\rangle^{\otimes n}) &= \frac{q_{1\to 0}^n \cos^2 n\varphi}{(1-q_{0\to 1})^n \sin^2 n\varphi + q_{1\to 0}^n \cos^2 n\varphi} \\ &= \left(1 + \left(\frac{1-q_{0\to 1}}{q_{1\to 0}}\right)^n \tan^2 n\varphi\right)^{-1}.\end{aligned} \tag{91}$$

This is an upper bound on the relative error rate over all possible weak values. It can be used to find a suitable n for a given error rate in the postselection, regardless of the magnitude of A_w. Moreover, when n is large (but $n\varphi \ll 1$), Eq. (91) can be simplified to

$$\lim_{A_w\to\infty} p(\text{error}||0\rangle^{\otimes n}) \approx \varphi^{-2} n^{-2} \left(\frac{1-q_{0\to 1}}{q_{1\to 0}}\right)^{-n}. \tag{92}$$

This implies that the upper bound of the relative error rate $p(\text{error}||0\rangle^{\otimes n})$ can also decay super-exponentially with n, which verifies the advantage of entanglement in suppressing the effect of readout errors.

Note that Eq. (92) does not contradict with (90). In (90), it is assumed that $n\varphi A_w \ll 1$, which requires A_w not to be too large, while in (92), we take the limit $A_w \to \infty$. Since the weak value amplification is a linear approximation theory which works in the regime $n\varphi A_w \ll 1$, Eq. (90) can be used in practice. And Eq. (92) is mainly to provide an upper bound for $p(\text{error}||0\rangle^{\otimes n})$.

B. Loss of correct postselection results

In this subsection, we turn to another important effect of readout errors: the loss of correct postselection results when $|0\rangle$ is misread as $|1\rangle$. Since the probability of postselecting the n entangled qubits in the state $|0\rangle^{\otimes n}$ is $p_0^{(n)}$, and among all postselection results $|0\rangle^{\otimes n}$ the proportion of correct states is $(1-q_{0\to 1})^n$, the probability of correct postselection results is

$$p(\text{correct}) = p_0^{(n)} (1-q_{0\to 1})^n. \tag{93}$$

The reduction factor $(1-q_{0\to 1})^n$ quantifies the loss of correct postselection results caused by readout errors in the postselection measurements. The loss rate for correct postselection is therefore

$$r_{\text{loss}} = 1 - (1-q_{0\to 1})^n. \tag{94}$$

If the readout error probability $q_{0\to 1}$ is small, so that $nq_{0\to 1} \ll 1$, then

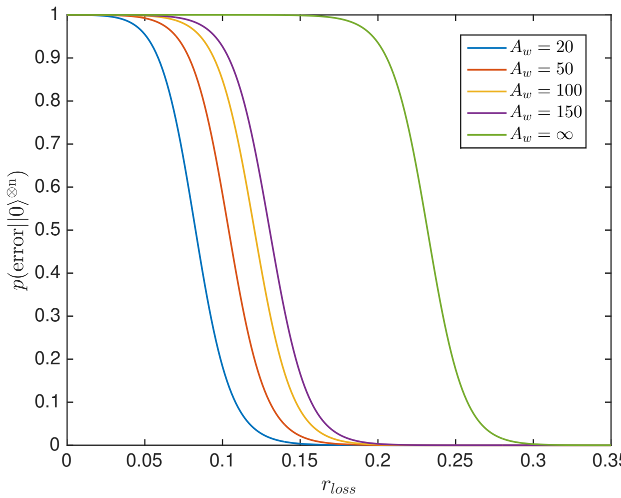
$$r_{\text{loss}} \approx nq_{0\to 1}. \tag{95}$$

Comparing (90) and (95), one finds that when n increases, the error rate falls but the loss rate grows, and vice versa. It implies a complementary relation between the error rate and the loss rate.

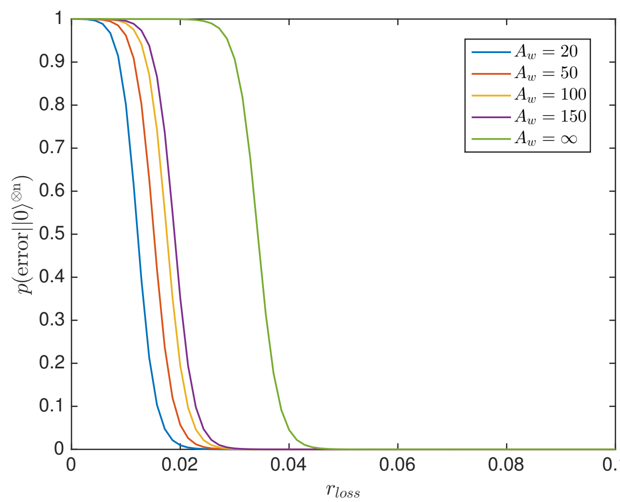
This relation for the limiting case $A_w \to \infty$ (which maximizes the relative error rate $p(\text{error}||0\rangle^{\otimes n})$) can be obtained in the following way. Suppose $\langle D|\hat{\sigma}_x|D\rangle = 0$ and $A_w \to \infty$, then by solving for n from (94) and plugging it into (91), we find that

$$\begin{aligned}&\lim_{A_w\to\infty} p(\text{error}||0\rangle^{\otimes n}) \\ &= \left(1 + (1-r_{\text{loss}}) q_{1\to 0}^{-\frac{\log(1-r_{\text{loss}})}{\log(1-q_{0\to 1})}} \tan^2 \frac{\varphi \log(1-r_{\text{loss}})}{\log(1-q_{0\to 1})}\right)^{-1}.\end{aligned} \tag{96}$$

It shows the complementary relation between the relative error rate and the loss rate due to readout errors when $A_w \to \infty$.



(a) $q_{1\to 0} = q_{0\to 1} = 0.05$



(b) $q_{1\to 0} = q_{0\to 1} = 0.01$

Figure 6. (Color online) Plot of the complementarity relation between the maximum relative error probability $p(\text{error}||0\rangle^{\otimes n})$ and the loss rate r_{loss} for different weak values A_w. The lines from left to right in each figure are for $A_w = 20, 50, 100, 150, \infty$, respectively. The r_{loss} axis in (6b) is rescaled.

In Fig. 6, the relative error rate $p(\text{error}||0\rangle^{\otimes n})$ is plotted versus the loss rate r_{loss} for different weak values. As before, $q_{0\to 1}$ and $q_{1\to 0}$ are assumed to be equal. The complementary relation between $p(\text{error}||0\rangle^{\otimes n})$ and r_{loss} can be explicitly observed there.

C. Influence on the measurement result

In the previous two subsections, we studied the two main effects of the readout errors: the relative error rate in total postselection results and the loss rate of correct postselection results. But in practical applications, what people care about is the final result of the measurement. So, how significantly do the readout errors affect the measurement result? And does the entanglement help to suppress their influence? It is important to make this question clear.

In this subsection, we investigate the modification of the average measurement result from the pointer qubit when readout errors exist in the postselection. We will obtain the modified average measurement result in the presence of readout errors, and show how entanglement can suppress the influence of readout errors on the measurement result.

Suppose we measure $\hat{\sigma}_x$ on the pointer qubit after postselecting the n system qubits, for example. According to $|D_{f,0}\rangle$ and $|D_{f,1}\rangle$ in (73) and (74), the average shift of the pointer is

$$\begin{aligned}\delta\langle\hat{\sigma}_x\rangle_0 &= -\frac{n\sin 2n\varphi \mathrm{Im}A_w(1-\langle\hat{\sigma}_x\rangle_D^2)}{\eta_0^{(n)}},\\ \delta\langle\hat{\sigma}_x\rangle_1 &= \frac{n\sin 2n\varphi \mathrm{Im}A_w(1-\langle\hat{\sigma}_x\rangle_D^2)}{\eta_1^{(n)}},\end{aligned}\tag{97}$$

for postselection states $|0\rangle^{\otimes n}$ and $|1\rangle^{\otimes n}$ respectively, where $\eta_0^{(n)}$ and $\eta_1^{(n)}$ are defined in Eq. (76). So, taking readout error into account, the real average result from the pointer qubit when each system qubit is postselected to $|0\rangle$ is

$$\begin{aligned}\overline{\delta\langle\hat{\sigma}_x\rangle} &= \frac{p_0^{(n)}(1-q_{0\to 1})^n\delta\langle\hat{\sigma}_x\rangle_0 + p_1^{(n)}q_{1\to 0}^n\delta\langle\hat{\sigma}_x\rangle_1}{p_0^{(n)}(1-q_{0\to 1})^n + p_1^{(n)}q_{1\to 0}^n}\\ &= -\gamma(n)\frac{n\sin 2n\varphi \mathrm{Im}A_w(1-\langle\hat{\sigma}_x\rangle_D^2)}{\eta_0^{(n)}},\end{aligned}\tag{98}$$

where

$$\gamma(n) = \frac{p_0^{(n)}((1-q_{0\to 1})^n - q_{1\to 0}^n)}{p_0^{(n)}(1-q_{0\to 1})^n + p_1^{(n)}q_{1\to 0}^n}.\tag{99}$$

Note that $\langle\hat{\sigma}_x^2\rangle_D = 1$. Therefore, $1-\langle\hat{\sigma}_x\rangle_D^2 = \mathrm{Var}(\hat{\sigma}_x)_D$, and thus,

$$\overline{\delta\langle\hat{\sigma}_x\rangle} = -\gamma(n)\frac{n\sin 2n\varphi \mathrm{Im}A_w \mathrm{Var}(\hat{\sigma}_x)_D}{\eta_0^{(n)}}.\tag{100}$$

When there is no readout error, the average result from the pointer qubit is $-n\sin 2n\varphi \mathrm{Im}A_w\mathrm{Var}(\hat{\sigma}_x)_D/\eta_0^{(n)}$, which is approximately $-2\varphi\mathrm{Im}A_w\mathrm{Var}(\hat{\sigma}_x)_D$ when $nA_w\varphi \ll 1$. Therefore, readout errors change the average measurement result of the pointer qubit by the overall factor $\gamma(n)$. Note that

$$\gamma(n) = 1 - \frac{p(\text{error}||0\rangle^{\otimes n})}{p_1^{(n)}},\tag{101}$$

from Eq. (87), which relates the correction factor $\gamma(n)$ to the relative error rate $p(\text{error}||0\rangle^{\otimes n})$.

According to Eq. (99), one can deduce that $|\gamma(n)| \le 1$, which means that readout errors reduce the weak value amplification, but never enhance it. This can be understood intuitively: the postselection results $|1\rangle^{\otimes n}$ do not give an amplification of φ as the results $|0\rangle^{\otimes n}$ do; so when they are mixed with the correct postselection results $|0\rangle^{\otimes n}$, the amplification factor will always fall.

Notably, since the relative error rate $p(\text{error}||0\rangle^{\otimes n})$ can decrease super-exponentially fast with the number of entangled qubits n according to (90), and $p_1^{(n)}$ is close to 1 in the weak value amplification, the correction factor $\gamma(n)$ can then be increased to 1 very efficiently with n. This is verified by the numerical results in Fig. 7, where the correction factor $\gamma(n)$ versus the number of entangled qubits, n, is plotted for different weak values (including $A_w = \infty$). The results again show how entanglement can significantly strengthen the postselected weak measurement against readout errors.

D. Influence on Fisher information

In the last subsection, it was shown that readout errors reduce the average measurement result by the factor γ (101), which is linear in the relative error rate $p(\text{error}||0\rangle^{\otimes n})$, and does not depend on the loss rate r_{loss}. This suggests that the error rate in the postselected events affects the result of the weak measurement, but the loss of correct postselected events does not. Since entanglement between the system qubits can dramatically decrease the error rate, it would be sufficient in this sense to use entanglement to suppress the effect of readout errors, regardless of the loss of correct postselected events, the rate of which can increase with the number of entangled qubits.

This is true for the average measurement result, because it does not depend on the size of the set of (correct) postselected events. The loss rate mainly affects the size of that set. However, the precision of estimating φ from the average measurement results does depend on the loss rate, since the estimation precision of a parameter generally relies on the size of the sample.

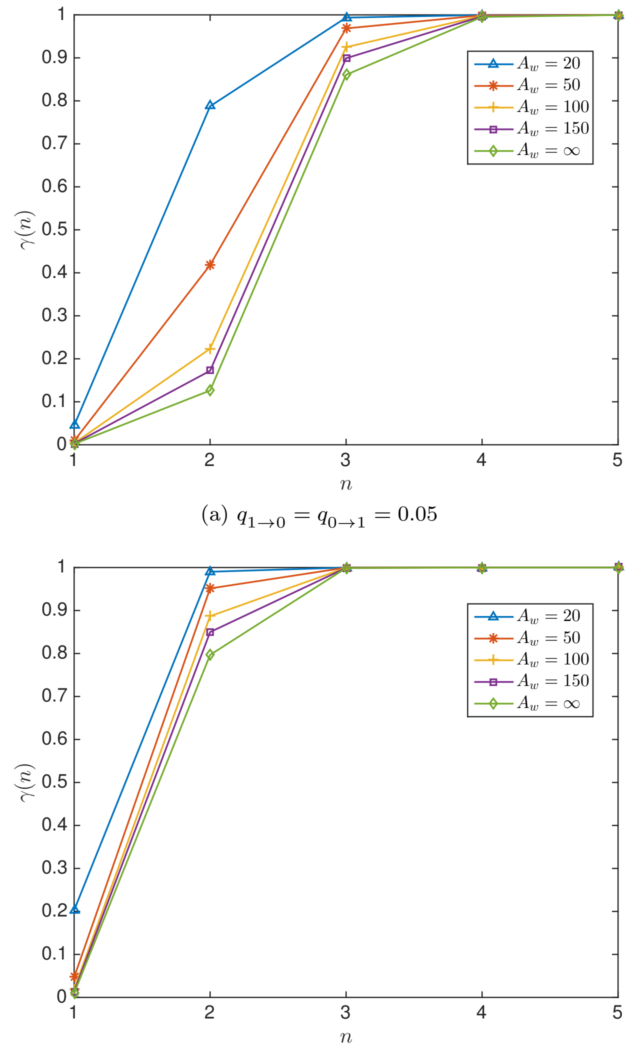
In this subsection, we study the effect of readout errors on the Fisher information of estimating φ in postselected weak measurements. We will show that the loss rate indeed affects the Fisher information. This implies that both the loss rate and the average error must be suppressed to maintain the precision of the measurement.

To compute the Fisher information, suppose we perform a POVM $\{\hat{E}_1, \cdots, \hat{E}_r\}$ on the pointer state after postselection. Then , when the n entangled qubits are postselected in the state $|0\rangle^{\otimes n}$, and the final pointer state is $|D_{f0}\rangle$, then the probability to observe the jth measurement outcome is

$$w_{0,j} = \langle D_f^0|\hat{E}_j|D_f^0\rangle. \tag{102}$$

Similarly, when the n entangled qubits are postselected in the state $|1\rangle^{\otimes n}$, the final pointer state is $|D_{f1}\rangle$, and the probability to observe the jth outcome is

$$w_{1,j} = \langle D_f^1|\hat{E}_j|D_f^1\rangle. \tag{103}$$



(a) $q_{1\to 0} = q_{0\to 1} = 0.05$

(b) $q_{1\to 0} = q_{0\to 1} = 0.01$

Figure 7. (Color online) This figure plots the correction factor $\gamma(n)$ of the measurement result versus the number of entangled qubits n for different weak values A_w. It can be seen that the correction factor can increase to 1 very fast with n, which shows the effectiveness of entanglement in protecting the weak value amplification against readout errors.

According to Eqs. (73) and (74), when $\varphi A_w \ll 1$, $|D_{f0}\rangle$ and $|D_{f1}\rangle$ can be approximated by

$$\begin{aligned} |D_{f0}\rangle &\approx \exp(-\mathrm{i}\varphi A_w \hat{\sigma}_x)|D\rangle, \\ |D_{f1}\rangle &\approx \exp\left(\mathrm{i}\varphi \frac{n^2}{A_w^*}\hat{\sigma}_x\right)|D\rangle, \end{aligned} \tag{104}$$

So the probabilities of obtaining the jth measurement outcome from pointer states $|D_{f0}\rangle$ and $|D_{f1}\rangle$ are, respectively,

$$\begin{aligned} w_{0,j} &= \frac{\langle D|\exp(\mathrm{i}\varphi A_w^* \hat{\sigma}_x)\hat{E}_j \exp(-\mathrm{i}\varphi A_w \hat{\sigma}_x)|D\rangle}{\langle D|\exp(2\varphi \mathrm{Im} A_w \hat{\sigma}_x)|D\rangle}, \\ w_{1,j} &= \frac{\langle D|\exp(-\mathrm{i}\frac{n^2}{A_w}\varphi\hat{\sigma}_x)\hat{E}_j \exp(\mathrm{i}\frac{n^2}{A_w^*}\varphi\hat{\sigma}_x)|D\rangle}{\langle D|\exp\left(-2\frac{n^2}{|A_w|^2}\varphi \mathrm{Im} A_w \hat{\sigma}_x\right)|D\rangle}. \end{aligned} \tag{105}$$

Therefore, the total probability of observing the jth outcome from the final pointer state is

$$h_j = p_0^{(n)} w_{0,j}(1-q_{0\to1})^n + p_1^{(n)} w_{1,j} q_{1\to0}^n, \tag{106}$$

where the probability of readout errors has been included.

Since $\varphi \ll 1$, we focus on the zeroth order of the Fisher information, i.e., $\varphi = 0$. The Fisher information we acquire from the pointer state by the POVM $\{\hat{E}_1, \cdots, \hat{E}_r\}$ is

$$I_\varphi = \sum_j \frac{(\partial_\varphi h_j)^2}{h_j}. \tag{107}$$

The term $\partial_\varphi h_j$ can be expanded as

$$\begin{aligned}\partial_\varphi h_j = &(w_{0,j}\partial_\varphi p_0^{(n)} + p_0^{(n)}\partial_\varphi w_{0,j})(1-q_{0\to1})^n \\ &+ (w_{1,j}\partial_\varphi p_1^{(n)} + p_1^{(n)}\partial_\varphi w_{1,j}) q_{1\to0}^n.\end{aligned} \tag{108}$$

From (105), we obtain that

$$\begin{aligned} w_{1,j}|_{\varphi=0} &= w_{0,j}|_{\varphi=0}, \\ \partial_\varphi w_{1,j}|_{\varphi=0} &= -\frac{n^2}{|A_w|^2}\partial_\varphi w_{0,j}|_{\varphi=0}, \end{aligned} \tag{109}$$

and Eqs. (75), (76) imply that

$$\begin{aligned} p_1^{(n)}|_{\varphi=0} &= \frac{|A_w|^2}{n^2} p_0^{(n)}|_{\varphi=0}, \\ \partial_\varphi p_1^{(n)}|_{\varphi=0} &= -\partial_\varphi p_0^{(n)}|_{\varphi=0}. \end{aligned} \tag{110}$$

Plugging these equations into (107), we get

$$\begin{aligned} I_\varphi = &\frac{n^2((1-q_{0\to1})^n - q_{1\to0}^n)^2}{n^2(1-q_{0\to1})^n + |A_w|^2 q_{1\to0}^n} \\ &\times \sum_j \frac{(w_{0,j}\partial_\varphi p_0^{(n)} + p_0^{(n)}\partial_\varphi w_{0,j})^2}{p_0^{(n)} w_{0,j}}. \end{aligned} \tag{111}$$

When there are no readout errors, $q_{1\to0} = q_{0\to1} = 0$, so the Fisher information is

$$I_{\varphi,0} = \sum_j \frac{(w_{0,j}\partial_\varphi p_0^{(n)} + p_0^{(n)}\partial_\varphi w_{0,j})^2}{p_0^{(n)} w_{0,j}}. \tag{112}$$

Therefore, the Fisher information modified by the readout errors can be written as

$$I_{\varphi,0} = f(n) I_{\varphi,0}, \tag{113}$$

where

$$f(n) = \frac{n^2((1-q_{0\to1})^n - q_{1\to0}^n)^2}{n^2(1-q_{0\to1})^n + |A_w|^2 q_{1\to0}^n}. \tag{114}$$

The factor $f(n)$ quantifies the effect of readout errors on the Fisher information. Note that $f(n) < 1$ when $q_{1\to0} \neq 0$, $q_{0\to1} \neq 0$, so readout errors always reduce the Fisher information of the weak measurement, and never increase it. Eq. (114) also shows the role of the number of entangled qubits n and the weak value A_w on the extent to which the readout errors can reduce the Fisher information.

Fig. 8 plots how the Fisher information changes as n increases. It shows that entanglement can recover some lost Fisher information by raising the average shift, but the Fisher information is still reduced by loss. As we will see in the next subsection, when entanglement is combined with a majority voting scheme, more Fisher information can be recovered.

Before concluding this subsection, we want to point out the relation between the modification factor $f(n)$ for Fisher information and the loss rate of correct postselected events r_{loss}. Generally, the probability of a single readout error is low, so $q_{1\to0}^n \to 0$ when n is not small. And according to Eq. (94), $(1-q_{0\to1})^n = 1-r_{\text{loss}}$. Hence, from (114) we immediately have

$$f(n) \approx 1 - r_{\text{loss}}. \tag{115}$$

This explicitly shows the relation between $f(n)$ and the loss rate r_{loss}, and verifies that the loss of correct postselected events does indeed cause a reduction in the Fisher information of the weak measurement.

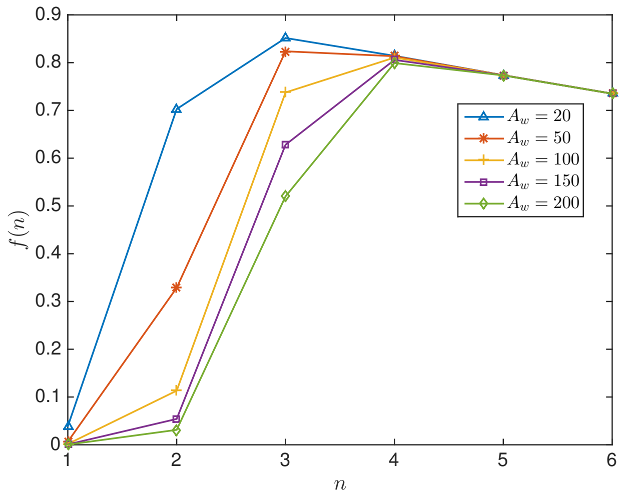
The relation (115) can also be understood in a more intuitive way: the Fisher information is proportional to the size of the sample that is used for parameter estimation, and the proportion of correct postselection results in the whole set of postselected events is $1 - r_{\text{loss}}$, so the Fisher information is reduced by r_{loss}, as indicated by (115). This suggests the necessity of suppressing the loss rate as well as the relative error rate.

E. Majority voting scheme for recovering Fisher information

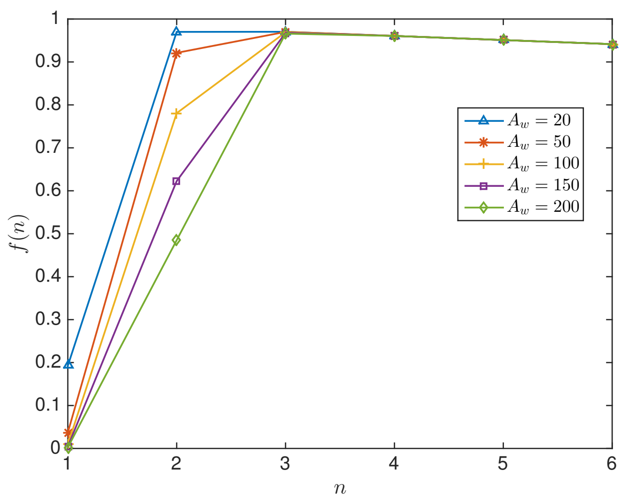
As the loss of correct postselection results can be detrimental to the Fisher information of the weak measurement, it is necessary to eliminate or suppress the loss. In this subsection, we introduce a majority voting scheme on the postselection results to decrease the loss rate of the correct postselection results and increase the effective Fisher information of the weak measurement.

The idea comes from a simple observation on Eq. (72): the true postselected states of the n system qubits are correlated, and should either be all $|0\rangle$ or all $|1\rangle$. So when readout errors occur, it is still possible to determine whether the postselection is successful or not with high probability from the majority of the observed states of the n qubits. If one observes more $|0\rangle$'s than $|1\rangle$'s, it is more likely that the n system qubits are postselected to $|0\rangle^{\otimes n}$. And vice versa. This is the majority voting scheme.

In this subsection, we will examine this scheme carefully, and show that it can effectively suppress the loss of the correct postselection results, and recover the lost Fisher information of the weak measurement.



(a) $q_{1\to 0} = q_{0\to 1} = 0.05$



(b) $q_{1\to 0} = q_{0\to 1} = 0.01$

Figure 8. (Color online) This figure plots the modification factor $f(n)$ of the Fisher information versus the number of entangled qubits n in the presence of readout errors, but without majority voting, for different weak values. When n is small, the Fisher information can increase with n, because the entanglement eliminates some errors in the postselection results and raises the proportion of correct postselected states that have higher Fisher information. However, when n becomes larger, the Fisher information starts to fall, since the loss rate of correct postselected states dramatically increases with n in this case.

Suppose at most k readout errors are allowed in the postselected results of a batch of n qubits (i.e. the number of observed $|0\rangle$'s or $|1\rangle$'s, whichever is lesser, is no more than some threshold k), and assume that the readout errors are independent of each other. The loss rate of the correct postselected states in this case becomes

$$r'_{\text{loss}} = 1 - \sum_{j=0}^{k} \binom{n}{j} (1 - q_{0\to 1})^{n-j} q_{1\to 0}^{j}, \tag{116}$$

which is obviously lower than Eq. (94), so more correct postselected events are retained by the majority voting scheme.

Now, let us study the Fisher information I_φ when majority voting is used. The total probability of postselecting the n qubits in the state $|0\rangle^{\otimes n}$ is

$$\begin{aligned} h_j^{(k)} = {} & p_0^{(n)} w_{0,j} \sum_{j=0}^{k} \binom{n}{j} (1 - q_{0\to 1})^{n-j} q_{1\to 0}^{j} \\ & + p_1^{(n)} w_{1,j} \sum_{j=0}^{k} \binom{n}{j} q_{1\to 0}^{n-j} (1 - q_{0\to 1})^{j}. \end{aligned} \tag{117}$$

Note that in the expansion of $\partial_\varphi h_j^{(k)}$ in this case, Eqs. (109) and (110) are unchanged. Therefore, the new Fisher information with the majority voting scheme can be derived by plugging the following replacement into Eq. (111):

$$\begin{aligned} (1 - q_{0\to 1})^{n} & \longrightarrow \sum_{j=0}^{k} \binom{n}{j} (1 - q_{0\to 1})^{n-j} q_{1\to 0}^{j}, \\ q_{1\to 0}^{n} & \longrightarrow \sum_{j=0}^{k} \binom{n}{j} q_{1\to 0}^{n-j} (1 - q_{0\to 1})^{j}. \end{aligned} \tag{118}$$

The result is

$$I_\varphi = f(n,k) I_{\varphi,0}, \tag{119}$$

where the factor $f(n,k)$ is

$$f(n,k) = \frac{n^2 \left(\sum_{j=0}^{k} \binom{n}{j} \left((1 - q_{0\to 1})^{n-j} q_{1\to 0}^{j} - q_{1\to 0}^{n-j} (1 - q_{0\to 1})^{j} \right) \right)^2}{\sum_{j=0}^{k} \binom{n}{j} \left(n^2 (1 - q_{0\to 1})^{n-j} q_{1\to 0}^{j} + |A_w|^2 q_{1\to 0}^{n-j} (1 - q_{0\to 1})^{j} \right)}, \tag{120}$$

and $I_{\varphi,0}$ is still the original Fisher information without readout error, the same as Eq. (112).

Just like the modification factor $f(n)$ in the case of readout errors without majority voting, the overall factor $f(n,k)$ in Eq. (119) determines the total change in the Fisher information due to readout errors in the presence of majority voting. When the probability of a readout error is sufficiently low, the $f(n,k)$ can be roughly approximated by

$$f(n,k) \approx 1 - r'_{\text{loss}}, \tag{121}$$

where r'_{loss} is given in (116), so it is only to be expected that part of the Fisher information will be recovered, since majority voting can retain some originally lost postselected events.

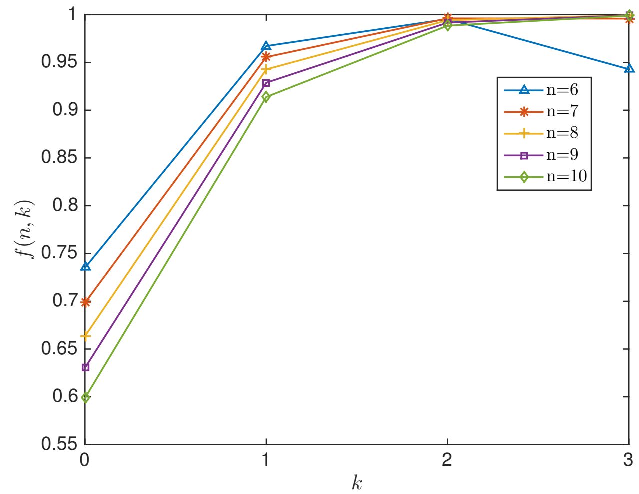
To see how efficiently the majority voting scheme can work for protecting the Fisher information against the readout errors, we study the factor $f(n,k)$ in detail numerically.

Fig. 9 plots the factor $f(n,k)$ versus k for different n. The line $k=0$ corresponds to the case without majority voting. The figure shows the modification factor $f(n,k)$ has a dramatic increase from $k=0$ to $k=1$ (and larger k), and can almost reach 1 with proper k. This implies that with the majority voting scheme, the loss of correct postselection results can be almost completely suppressed, and nearly all of the lost Fisher information can be recovered.

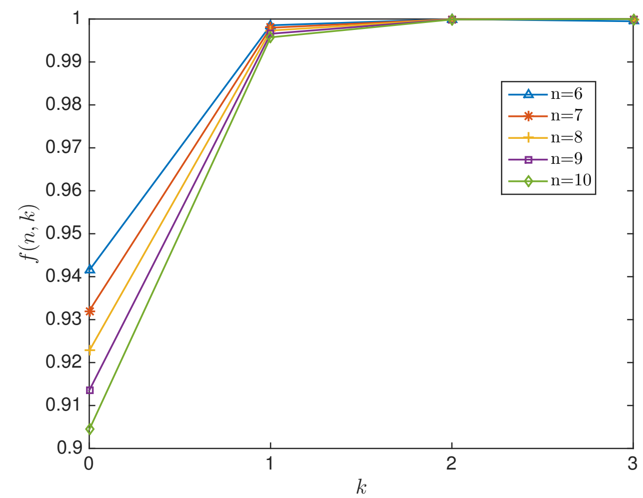
This is a remarkable result. By contrast, when no majority voting scheme is used, the lost Fisher information can only be partially recovered by the entanglement, as indicated by Fig. 8. This verifies the effectiveness of the majority voting scheme in protecting Fisher information against readout errors.

A notable point in Fig. 9a is that for $n=6$, when $k=3$, there is a small drop in $f(n,k)$. This is because when n is even and $k=n/2$, if the probability of readout errors is not small, a large fraction of the failed postselections will be identified as successful ones, and the failed postselections contain much lower Fisher information than the successful ones. So the Fisher will fall in this case. However, if the probability of readout errors is sufficiently low, the fraction of misidentified postselections will be very small, then the Fisher information will not drop. Fig. 9b shows the latter case.

It is also worth mentioning that according to Eq. (119), the effect of readout errors with the majority voting scheme on the Fisher information is an overall reduction by the factor $f(n,k)$, so when the original Fisher information $I_{\varphi,0}$ is maximized, the reduced Fisher information I_{φ} is also maximized. This implies that the optimal measurement to maximize the original Fisher information will still be the optimal when readout errors exist and the majority voting scheme is used. So the optimal measurement on the pointer qubit does not need to change in the presence of readout errors. This may be convenient for practical applications.



(a) $q_{1\to0}=q_{0\to1}=0.05$



(b) $q_{1\to0}=q_{0\to1}=0.01$

Figure 9. (Color online) This figure plots the modification factor $f(n,k)$ of the Fisher information versus the allowed number of readout errors k for different n when the majority voting scheme is used. The number of allowed readout errors, k, is from 0 to 3. The weak value is $A_w=30$. The figure shows that when k climbs to half of n, $f(n,k)$ can be very close to 1. It implies that with a proper number of readout errors allowed in postselecting the n qubits, the majority voting strategy can recover almost all the Fisher information lost by the readout errors.

F. Effect of majority voting scheme on the measurement result

In the last subsection, we showed that the majority voting scheme can efficiently recover almost all of the Fisher information lost by readout errors. A separate question is how the measurement result is affected by the majority voting scheme. In this subsection, we will investigate this problem in detail. We still measure $\hat{\sigma}_x$ on the pointer qubit after postselecting the n system qubits, similar to Sec. VII C.

According to Eq. (97), if we allow at most k errors in the postselection result of the n system qubits, the real average result from the pointer qubit after the postselection is

$$\begin{aligned}
&\overline{\delta\langle\hat{\sigma}_x\rangle} \\
&=\Bigg(\sum_{j=0}^{k}\binom{n}{j}\Big(p_0^{(n)}(1-q_{0\to1})^{n-j}q_{1\to0}^{j}\delta\langle\hat{\sigma}_x\rangle_0 \\
&\quad +p_1^{(n)}(1-q_{0\to1})^{j}q_{1\to0}^{n-j}\delta\langle\hat{\sigma}_x\rangle_1\Big)\Bigg)\Big/ \\
&\quad\Bigg(\sum_{j=0}^{k}\binom{n}{j}\Big(p_0^{(n)}(1-q_{0\to1})^{n-j}q_{1\to0}^{j} \\
&\quad + p_1^{(n)}(1-q_{0\to1})^{j}q_{1\to0}^{n-j}\Big)\Bigg) \\
&=-\gamma(n,k)\frac{n\sin 2n\varphi \mathrm{Im}A_w \mathrm{Var}(\hat{\sigma}_x)_D}{\eta_0^{(n)}},
\end{aligned} \tag{122}$$

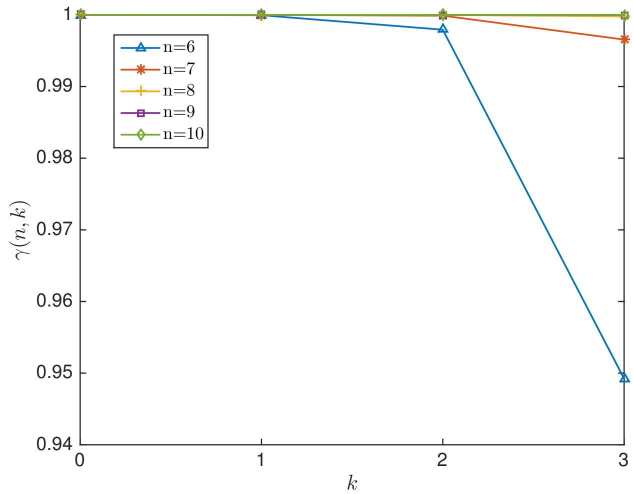
where we have used $(1-\langle\hat{\sigma}_x\rangle_D^2)=\mathrm{Var}(\hat{\sigma}_x)_D$, and

$$\gamma(n,k)= \frac{p_0^{(n)}\sum_{j=0}^{k}\binom{n}{j}\Big((1-q_{0\to1})^{n-j}q_{1\to0}^{j}-(1-q_{0\to1})^{j}q_{1\to0}^{n-j}\Big)}{\sum_{j=0}^{k}\binom{n}{j}\Big(p_0^{(n)}(1-q_{0\to1})^{n-j}q_{1\to0}^{j}+p_1^{(n)}(1-q_{0\to1})^{j}q_{1\to0}^{n-j}\Big)}. \tag{123}$$

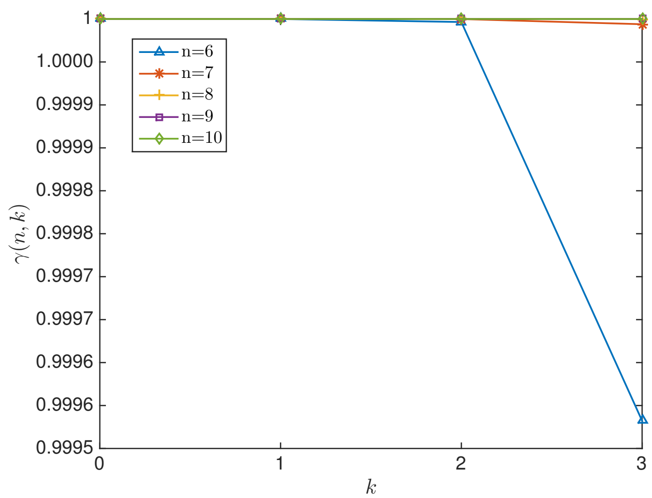
Similar to the case without majority voting, $\gamma(n,k)$ is the correction factor of the average measurement result, using the majority voting scheme with at most k errors allowed. From Eq. (123) one can deduce that $|\gamma(n,k)|\leq 1$, so the readout errors still reduce the weak value amplification even when majority voting is used.

Fig. 10 plots the factor $\gamma(n,k)$ versus k for different n. When k increases, the amplification of the measurement result has a small drop. The reason is similar to that for the drop in the Fisher information in Fig. 9a. That is, when k increases, more errors are allowed by the majority voting strategy, and erroneous postselections correspond to much lower weak values than correct postselections, so the amplification factor is reduced.

The results of this section and the last section suggest that there is a balance between the number of entangled qubits n and the allowed number of readout errors k in the majority voting scheme. On the one hand, to effectively restore the lost Fisher information, k should not be too small; otherwise, the majority voting scheme cannot recover most of the lost successful postselections, and a considerable part of Fisher information will still be discarded. On the other hand, if k is too large, there will be too many wrong postselection results mixed into the correction postselections, which will decrease both the Fisher information and the weak value amplification factor. Therefore, for a given number of entangled qubits n, one needs to find a suitable number of allowed readout errors k, so that the effects of readout errors can be effectively suppressed.



(a) $q_{1\to0}=q_{0\to1}=0.05$



(b) $q_{1\to0}=q_{0\to1}=0.01$

Figure 10. (Color online) This figure plots the correction factor $\gamma(n,k)$ of the measurement result versus the allowed number of readout errors k for different n, when the majority voting scheme is employed. The weak value is $A_w=30$. The lines for $n=8,9,10$ almost overlap, since they are very close to each other. The figure shows that when k increases, a small drop in the measurement result may occur, due to more errors being induced by the majority voting scheme.

VIII. SUMMARY

In this paper, we studied the optimization of postselected weak measurements to improve the performance of weak value amplification. This problem is approached in two ways: one is to maximize the postselection probability with a fixed weak value, which aims to improve the usage of resources; the other is to maximize the weak value for a given postselection probability, which aims to enhance the amplification ability of weak measurements.

We found that both of these can be significantly increased by using entangled systems, which results in that the Fisher information of the measurement can also be increased, and can approximately saturate the Heisenberg limit. Based on this, we proposed a protocol for entanglement-assisted weak measurement. We provided the optimal choice of initial state and postselection of the system for this protocol, and illustrated it by a qubit example with simple quantum circuits.

Furthermore, we considered the influence of readout errors on the protocol. Readout errors are more harmful to postselected weak measurements than to other quantum measurements, since even a small rate of readout errors can give rise to severe disturbance in a postselected weak measurement when the postselection probability is low. So it is particularly necessary to consider the effect of readout errors in postselected weak measurements.

There are two major problems resulting from readout errors. One is that an error will occur in the measurement result; the other is that the Fisher information of the measurement will decrease. We found that entanglement between the systems can eliminate the error in the measurement result very efficiently, so the first problem can be solved. Moreover, entanglement can also retrieve some of the lost Fisher information. To further suppress the loss of Fisher information, we introduced a majority voting strategy, and showed that with this strategy, almost no Fisher information will be lost.

Postselected weak measurement is a useful scheme to measure tiny physical effects, and how to exploit quantum resources to overcome its low efficiency and improve the sensitivity is of great interest in practical applications. It is worth mentioning that recently squeezing was also found useful in increasing the SNR of postselected weak measurement [64]. We hope that our work will help to deepen the understanding of this innovative measurement protocol, and extend it to broader applications.

ACKNOWLEDGMENTS

We thank Justin Dressel for much useful discussion. 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] T. Koike and S. Tanaka, Phys. Rev. A **84**, 062106 (2011).
[3] Y. Susa, Y. Shikano, and A. Hosoya, Phys. Rev. A **85**, 052110 (2012).
[4] A. Di Lorenzo, Annals of Physics **345**, 178 (2014).
[5] S. Pang, T. A. Brun, S. Wu, and Z.-B. Chen, Phys. Rev. A **90**, 012108 (2014).
[6] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, Phys. Rev. D **40**, 2112 (1989).
[7] R. Jozsa, Phys. Rev. A **76**, 044103 (2007).
[8] G. J. Pryde, J. L. O'Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, Phys. Rev. Lett. **94**, 220405 (2005).
[9] O. Hosten and P. Kwiat, Science **319**, 787 (2008).
[10] Y. Gorodetski, K. Y. Bliokh, B. Stein, C. Genet, N. Shitrit, V. Kleiner, E. Hasman, and T. W. Ebbesen, Phys. Rev. Lett. **109**, 013901 (2012).
[11] X. Zhou, Z. Xiao, H. Luo, and S. Wen, Phys. Rev. A **85**, 043809 (2012).
[12] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Phys. Rev. Lett. **66**, 1107 (1991).
[13] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Phys. Rev. Lett. **102**, 173601 (2009).
[14] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A **80**, 041803 (2009).
[15] M. Pfeifer and P. Fischer, Opt. Express **19**, 16508 (2011).
[16] M. D. Turner, C. A. Hagedorn, S. Schlamminger, and J. H. Gundlach, Opt. Lett. **36**, 1479 (2011).
[17] G. I. Viza, J. Martínez-Rincón, G. B. Alves, A. N. Jordan, and J. C. Howell, arXiv:1410.8461 [quant-ph] (2014).
[18] S. Goswami, M. Pal, A. Nandi, P. K. Panigrahi, and N. Ghosh, arXiv:1407.6109 [physics] (2014).
[19] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, Phys. Rev. A **82**, 011802 (2010).
[20] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A **82**, 063822 (2010).
[21] X.-Y. Xu, Y. Kedem, K. Sun, L. Vaidman, C.-F. Li, and G.-C. Guo, Phys. Rev. Lett. **111**, 033604 (2013).
[22] P. Egan and J. A. Stone, Opt. Lett. **37**, 4991 (2012).
[23] G. Strübi and C. Bruder, Phys. Rev. Lett. **110**, 083605 (2013).
[24] G. I. Viza, J. Martínez-Rincón, G. A. Howland, H. Frostig, I. Shomroni, B. Dayan, and J. C. Howell, Opt. Lett. **38**, 2949 (2013).
[25] M. Mirhosseini, G. Viza, O. S. Magaña Loaiza, M. Malik, J. C. Howell, and R. W. Boyd, arXiv:1412.3019 [physics, physics:quant-ph] (2014).
[26] N. Brunner and C. Simon, Phys. Rev. Lett. **105**, 010405 (2010).
[27] A. Feizpour, X. Xing, and A. M. Steinberg, Phys. Rev. Lett. **107**, 133603 (2011).
[28] C.-F. Li, X.-Y. Xu, J.-S. Tang, J.-S. Xu, and G.-C. Guo, Phys. Rev. A **83**, 044102 (2011).
[29] O. Zilberberg, A. Romito, and Y. Gefen, Phys. Rev. Lett. **106**, 080405 (2011).
[30] J. B. Götte and M. R. Dennis, New J. Phys. **14**, 073016 (2012).
[31] A. Nishizawa, K. Nakamura, and M.-K. Fujimoto, Phys. Rev. A **85**, 062108 (2012).
[32] S. Wu and M. Żukowski, Phys. Rev. Lett. **108**, 080403 (2012).
[33] J. Dressel, K. Lyons, A. N. Jordan, T. M. Graham, and P. G. Kwiat, Phys. Rev. A **88**, 023821 (2013).
[34] A. Hayat, A. Feizpour, and A. M. Steinberg, Phys. Rev. A **88**, 062301 (2013).
[35] L. Zhou, Y. Turek, C. P. Sun, and F. Nori, Phys. Rev. A **88**, 053815 (2013).
[36] S. Huang and G. S. Agarwal, arXiv:1501.02359 [quant-ph] (2015).
[37] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve, and A. N. Korotkov, Nat Phys **6**,

442 (2010).
[38] 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).
[39] P. Campagne-Ibarcq, L. Bretheau, E. Flurin, A. Auffèves, F. Mallet, and B. Huard, Phys. Rev. Lett. **112**, 180402 (2014).
[40] D. Lu, A. Brodutch, J. Li, H. Li, and R. Laflamme, New J. Phys. **16**, 053015 (2014).
[41] Y. Shikano, arXiv:1110.5055 [math-ph, physics:physics, physics:quant-ph] (2011).
[42] A. G. Kofman, S. Ashhab, and F. Nori, Physics Reports **520**, 43 (2012).
[43] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Rev. Mod. Phys. **86**, 307 (2014).
[44] J. Dressel, Phys. Rev. A **91**, 032116 (2015).
[45] X. Zhu, Y. Zhang, S. Pang, C. Qiao, Q. Liu, and S. Wu, Phys. Rev. A **84**, 052111 (2011).
[46] G. C. Knee, G. A. D. Briggs, S. C. Benjamin, and E. M. Gauger, Phys. Rev. A **87**, 012115 (2013).
[47] A. Nishizawa, arXiv:1504.04777 [quant-ph] (2015).
[48] C. Ferrie and J. Combes, Phys. Rev. Lett. **112**, 040406 (2014).
[49] J. Combes, C. Ferrie, Z. Jiang, and C. M. Caves, Phys. Rev. A **89**, 052117 (2014).
[50] S. Tanaka and N. Yamamoto, Phys. Rev. A **88**, 042116 (2013).
[51] L. Zhang, A. Datta, and I. A. Walmsley, Phys. Rev. Lett. **114**, 210801 (2015).
[52] A. N. Jordan, J. Tollaksen, J. E. Troupe, J. Dressel, and Y. Aharonov, Quantum Stud.: Math. Found. **2**, 5 (2015).
[53] A. N. Jordan, J. Martínez-Rincón, and J. C. Howell, Phys. Rev. X **4**, 011031 (2014).
[54] G. C. Knee and E. M. Gauger, Phys. Rev. X **4**, 011032 (2014).
[55] Y. Kedem, Phys. Rev. A **85**, 060102 (2012).
[56] G. C. Knee, J. Combes, C. Ferrie, and E. M. Gauger, arXiv:1410.6252 [physics, physics:quant-ph] (2014).
[57] S. Pang, J. Dressel, and T. A. Brun, Phys. Rev. Lett. **113**, 030401 (2014).
[58] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[59] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).
[60] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[61] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Annals of Physics **247**, 135 (1996).
[62] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge; New York, 2000).
[63] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, Cambridge, United Kingdom; New York, 2013).
[64] S. Pang and T. A. Brun, arXiv:1409.2567 [quant-ph] (2014).