Physical description of statistical hypothesis testing for a weak-value-amplification experiment using a birefringent crystal

Yuki Susa*
Department of Physics, Tokyo Institute of Technology, Tokyo 152-8551, Japan
(Received 27 May 2015; published 20 August 2015)

We investigate the weak measurement experiment demonstrated by Ritchie *et al.* [N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Phys. Rev. Lett. **66**, 1107 (1991)] from the viewpoint of the statistical hypothesis testing for the weak-value amplification proposed by Susa and Tanaka [Y. Susa and S. Tanaka, Phys. Rev. A **92**, 012112 (2015)]. We conclude that the weak-value amplification is a better method to determine whether the crystal used in the experiment is birefringent than the measurement without postselection, when the angles of two polarizers are almost orthogonal. This result gives a physical description and intuition of the hypothesis testing and supports the experimental usefulness of the weak-value amplification.

DOI: 10.1103/PhysRevA.92.022118 PACS number(s): 03.65.Ta, 03.67.−a, 42.25.Ja, 42.25.Lc

I. INTRODUCTION

In 1988, the weak measurement was proposed by Aharonov *et al.* as an indirect quantum measurement with the postselection of the measured system [1]. Many theoretical and experimental studies have been done for the weak measurement in recent years [2]. Some researchers focused on the usefulness of the weak measurement as a technique for a high-precision measurement [3]. From the weak measurement, we can extract a weak value, which can be outside the range of the eigenvalues [4] or even complex. The weak value is defined by

$$\langle \hat{A} \rangle_w = \frac{\langle f|\hat{A}|i\rangle}{\langle f|i\rangle}, \tag{1}$$

where $\hat{A}$ is an observable and $|i\rangle$ and $|f\rangle$ are the initial and the final states of the measured system, respectively. The weak measurement magnifies the output more than the one given by an ordinary projective measurement. The weak value appears as a shift of the probe wave function induced by an interaction between the measured system and the measuring probe after postselecting the final state of the measured system [5].

Actually, several experiments confirmed the weak-value amplification (WVA) effect [6–8]. Some theoretical papers have shown that WVA is robust against systematic or technical error [9,10]. On the other hand, there is a statistical argument that WVA has a disadvantage in the parameter estimation for the interaction strength, because the postselection makes the number of detectable data small [11–15]. The countercriticism also arose that the data loss by postselection is not critical in practical cases [16].

Reference [17] shows that WVA can be more significant for interaction detection than the measurement without postselection when the weak value is outside the eigenvalue range of the measured system observable. The WVA is more likely than the ordinary measurement to correctly indicate when the interaction indeed exists. This fact is derived by hypothesis testing [18], which is one of the methods of statistical inference. There it is supposed that the interaction is present if the measured value sufficiently deviates from the initial probe fluctuation. The decision method gives a uniformly most powerful unbiased (UMPU) test, i.e., a statistically good testing.

In this work we demonstrate the testing for WVA [17] by comparing the power defined in the statistical hypothesis testing in two measurements, the weak measurement and the ordinary one without postselection, for a particular experimental setup. We pick up the classic experiment using a birefringent crystal and two polarizers demonstrated by Ritchie *et al.* [6] for example. This experiment was originally designed for the measurement of the weak value. Here we look at the same experiment from a different angle. We regard it as a testing problem to distinguish whether the crystal is birefringent or not. Thus the statistical power in this experimental setup is given as the probability to determine exactly when the crystal is indeed birefringent. This experiment gives intensity distributions as the results of the weak measurement including the case that the weak-coupling approximation [1] does not hold and the ordinary one with the birefringent crystal. These experimental results clearly show the physical intuition of the testing and the advantage of WVA by comparing the powers as presented in Ref. [17]. We conclude that the angle of the polarizers that give the weak value is the only factor in determining the case that the weak measurement is superior. For other more recent WVA experiments, see the discussion in Sec. IV.

This paper is structured as follows. In Sec. II we introduce the weak measurement of the experiment presented by Ritchie *et al.* In Sec. III we implement the hypothesis testing to decide whether the crystal used in the experiment is birefringent or not. In Sec. IV we summarize this paper and discuss the issue of the number of detectable data in the actual case. Some equations are provided in the Appendix for convenience.

II. REVIEW OF BIREFRINGENCE EXPERIMENT FOR WEAK MEASUREMENT

Here we briefly review the optical setup of the experiment by Ritchie *et al.* [6]. In this experiment the position y of the laser beam is the measuring probe and the polarization is regarded as the measured system. We use the Gaussian shape

*susa@th.phys.titech.ac.jp

for the laser beam profile given by

$$\psi(y) = \langle y|\psi\rangle = \left(\frac{1}{2\pi w_0^2}\right)^{1/4} e^{-y^2/4w_0^2}. \tag{2}$$

The waist of the beam used in the experiment [6] is $w_0 = 55\ \mu$m. The tunings of the two polarizers play the roles of the pre- and postselection of the initial and final states in the measured system, which are

$$|i\rangle = \cos\alpha|H\rangle + \sin\alpha|V\rangle, \tag{3}$$

$$|f\rangle = \cos\beta|H\rangle + \sin\beta|V\rangle, \tag{4}$$

where $|H\rangle$ and $|V\rangle$ denote the horizontal and vertical polarization states, respectively, and α and β represent the angles of the first and second polarizers, respectively.

The experimental setup is displayed in Fig. 1. Column (i) in Fig. 1 shows the initial state of the polarization and the probe distribution. Passing through the first polarizer, the photons are injected into the birefringent crystal, which gives two different refraction factors depending on the polarization of the injected photons. The injected beam is spatially separated into two beams with different polarizations, one of which is called an ordinary ray and the other is an extraordinary ray. Then it gives the correlation between the position of the beam and the polarization. The refraction of the photons is described by the von Neumann–type Hamiltonian as

$$\hat{H} = g\delta(t)\hat{A} \otimes \hat{p}_y, \tag{5}$$

where g is the interaction strength, $\hat{A}$ is the observable given below, and $\hat{p}_y$ is the momentum operator conjugate to the position y of the measuring probe. Because the refraction factor depends on the polarization, $\hat{A}$ is given by

$$\hat{A} = \lambda_H|H\rangle\langle H| + \lambda_V|V\rangle\langle V|, \tag{6}$$

where $\lambda_{H,V} > 0$ are the eigenvalues of $\hat{A}$. We note that the crystal used in the experiment is a quartz plate, the refraction factors of which are $n_e = 1.551\,65$ for the extraordinary ray and $n_o = 1.542\,61$ for the ordinary ray when the wavelength of the injected laser is 633 nm as quoted by Ref. [6]. We also note that if the crystal is not birefringent, $\lambda_H = \lambda_V$.

The distribution function of the measuring probe after refraction is calculated as

$$\begin{aligned} &f_{\rm nps}(y|\lambda_H,\lambda_V) \\ &\quad = |\langle y|e^{-ig\hat{A}\otimes\hat{p}_y}|\psi\rangle|i\rangle|^2 \\ &\quad = \frac{1}{\sqrt{2\pi w_0^2}}\left(\cos^2\alpha\, e^{-(y-g\lambda_H)^2/2w_0^2} + \sin^2\alpha\, e^{-(y-g\lambda_V)^2/2w_0^2}\right). \end{aligned} \tag{7}$$

This output is obtained by the ordinary measurement without postselection. As we can see, it is composed of the two Gaussian distributions.

According to Ref. [6], the crystal thickness is $d = 331\ \mu$m and the angle of incidence is $\theta \approx 30^\circ$. Using Snell's law, we can associate $g\lambda_H$ and $g\lambda_V$ with the position shifts by the refractions as

$$g\lambda_H = d\frac{\sin(\theta-\theta_e)}{\cos\theta_e} \approx 67.92\ \mu\text{m}, \tag{8}$$

$$g\lambda_V = d\frac{\sin(\theta-\theta_o)}{\cos\theta_o} \approx 67.28\ \mu\text{m}, \tag{9}$$

where $\sin\theta_e = \sin\theta/n_e$ and $\sin\theta_o = \sin\theta/n_o$. Here we have regarded the horizontal and vertical polarized beams as extraordinary and ordinary rays, respectively. We see that the difference of 0.64 μm between $g\lambda_H$ and $g\lambda_V$ in the birefringence is much smaller than the beam waist $w_0 = 55\ \mu$m. Because the two beams almost overlap, we would observe a single-peak distribution, the median of which is $y = g(\lambda_H + \lambda_V)/2$, as the final probe distribution obtained by the photodetector. Then it is difficult to distinguish whether or not the crystal is birefringent [Fig. 1(c)(ii)]. We note that the polarization of the overlapped region is the same as the initial polarization, while the polarization of the nonoverlapped region is as shown in Fig. 1(b)(ii).

The probability distribution will be changed from $f_{\rm nps}(y)$ due to postselection by the second polarizer. The distribution

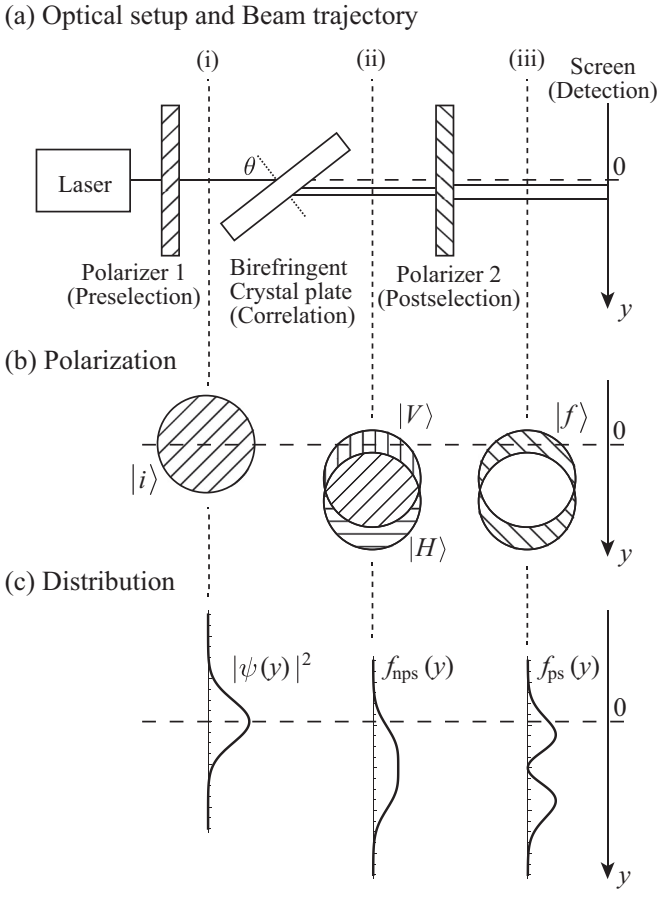

FIG. 1. (a) Sketch of the optical setup. The tilted birefringent crystal plates is set between the two polarizers, the angles of which are tuned almost orthogonal. On the screen, we observe the beam position y and we decide whether or not the crystal is birefringent. Also shown are (b) the beam polarization and (c) the probe distribution, in each stage. For the illustration, the angles of the polarizers are $\alpha = \pi/4$ and $\beta = 3\pi/4$.

function obtained by the weak measurement becomes

$$f_{\rm ps}(y|\lambda_H,\lambda_V) = \frac{|\langle y|\langle f|e^{-ig\hat{A}\otimes\hat{p}_y}|i\rangle|\psi\rangle|^2}{|\langle f|e^{-ig\hat{A}\otimes\hat{p}_y}|i\rangle|\psi\rangle|^2} = \frac{\cos^2\alpha\cos^2\beta e^{-(y-g\lambda_H)^2/2w_0^2}+\sin^2\alpha\sin^2\beta e^{-(y-g\lambda_V)^2/2w_0^2}+(1/2)\sin 2\alpha\sin 2\beta e^{-(1/2w_0^2)\{y-g[(\lambda_H+\lambda_V)/2]\}^2-(g^2/2w_0^2)[(\lambda_H-\lambda_V)/2]^2}}{\sqrt{2\pi w_0^2}[\cos^2\alpha\cos^2\beta+\sin^2\alpha\sin^2\beta+(1/2)\sin 2\alpha\sin 2\beta\, e^{-(g^2/2w_0^2)[(\lambda_H-\lambda_V)/2]^2}]}. \tag{10}$$

The third term of the numerator represents the interference. We remark that the coefficient $\sin 2\alpha \sin 2\beta$ can be negative when the angles of the two polarizers are nearly orthogonal. Its negativity tends to depress the final probe distribution in the central region and separate it into the two-peak distribution, because the polarization of the central part just after the crystal is orthogonal to the one of the second polarizer, which is used as the postselection [Fig. 1(b)(ii)]. The weak value becomes

$$\langle\hat{A}\rangle_w = \left(\frac{\lambda_H-\lambda_V}{2}\right)\frac{\cos(\alpha+\beta)}{\cos(\alpha-\beta)}+\frac{\lambda_H+\lambda_V}{2}. \tag{11}$$

When $\lambda_H \neq \lambda_V$, we can see that the weak value will be large for the almost orthogonal pair of α and β. Then we could observe the amplified peak-to-peak distance by using WVA [Fig. 1(c)(iii)] to conclude that the crystal is birefringent. However, the disadvantage of the postselection is the decrease of the entire intensity.

III. HYPOTHESIS TESTING FOR BIREFRINGENCE

As an application we consider the statistical hypothesis testing proposed in Ref. [17] in the experiment that uses the birefringent crystal as explained in the previous section. For the testing problem to determine whether the crystal is birefringent or not, we take the following hypotheses: the null hypothesis H_0, in which the crystal is not birefringent (i.e., $\lambda_H = \lambda_V$), and the alternative hypothesis H_1, in which the crystal is birefringent (i.e., $\lambda_H \neq \lambda_V$). We note that by the interaction, the refraction occurs for both hypotheses, which somewhat differs from the previous work [17].

We compare the testing power given by the weak measurement and the one given by the measurement without postselection. The testing power is defined as

$$b(\lambda_H,\lambda_V) := \int d(y) f(y|\lambda_H,\lambda_V)dy, \tag{12}$$

where $d(y)$ is a decision function, which is a mathematical expression for a decision criterion. The function $d(y)$ takes a binary value of 0 or 1. The 0 indicates that the null hypothesis is accepted and the 1 represents that the alternative one is accepted. The power (12) with $\lambda_H \neq \lambda_V$ indicates the probability to correctly judge that the alternative hypothesis is actually true.

In the previous research [17] that treats the testing problem to determine whether the interaction is present ($g \neq 0$) or absent ($g = 0$) between the measured system and the measuring probe, the proposed decision function

$$d(y) = \begin{cases} 0 & \text{for } |y|/w_0 < c \\ 1 & \text{for } |y|/w_0 \geqslant c \end{cases} \tag{13}$$

works well. Here c is a critical point indicating the threshold beyond which the null hypothesis is rejected. The physical interpretation of the decision function (13) is that the interaction would be present if the observed position y is outside the initial laser beam waist w_0.

However, this decision function does not suit the present birefringence testing problem as it is because the refraction could make the beam position y shifted outside the initial beam waist w_0, although the null hypothesis is actually true. Then, for the proper testing, we have to adjust the final probe wave function by shifting $\lambda_+ := (\lambda_H+\lambda_V)/2$. When the null hypothesis is true, for example, $g\lambda_+$ is the medium of a single Gaussian distribution for the final probe state. On the other hand, when the alternative hypothesis is true, $g\lambda_+$ coincides with the mean of the two peaks of the final probe distribution. We can grasp the value of $g\lambda_+$ by the preparatory experiment without postselection to just monitor the refraction by the crystal. Thus, the final probe distribution can be adjusted by a translation such as $y \to y + g\lambda_+$.

We remark that this adjustment can be described as the unitary operator $\exp[ig\lambda_+\hat{I}\otimes\hat{p}_y]$, where $\hat{I}$ is the identity operator of the measured system. Then the total unitary operator combining the two unitary operators given by the Hamiltonian (5) and by the adjustment becomes

$$\hat{U}_t = \exp[ig\lambda_+\hat{I}\otimes\hat{p}_y]\exp[-ig\hat{A}\otimes\hat{p}_y] = \exp[-ig(\hat{A}-\lambda_+\hat{I})\otimes\hat{p}_y] = \exp[-ig\hat{A}_t\otimes\hat{p}_y]. \tag{14}$$

Here we have introduced the total observable $\hat{A}_t := \lambda_-(|H\rangle\langle H| - |V\rangle\langle V|)$ and its eigenvalue $\lambda_- := (\lambda_H-\lambda_V)/2$ for convenience. Now we can rewrite our two hypotheses as $H_0 : \lambda_- = 0$ and $H_1 : \lambda_- \neq 0$ and the total weak value as

$$\langle\hat{A}_t\rangle_w = \frac{\langle f|\hat{A}_t|i\rangle}{\langle f|i\rangle} = \lambda_-\frac{\cos(\alpha+\beta)}{\cos(\alpha-\beta)}. \tag{15}$$

To calculate the testing powers of each measurement, we introduce (in the Appendix) the adjusted distributions: Eq. (A1) for the ordinary measurement without postselection and Eq. (A2) for the weak measurement. In the previous study (see Sec. III B in Ref. [17]) we showed that the decision function (13) becomes the UMPU test for the previous testing problem of $H_0 : g = 0$ and $H_1 : g \neq 0$, while the present testing problem is $H_0 : \lambda_- = 0$ and $H_1 : \lambda_- \neq 0$. Noting that

TABLE I. Condition $C(\alpha,\beta)$ and the weak value are displayed with the assigned values of α and β used in the actual experiment [6]. In each case, $\alpha = \pi/4$.

Case	β	$C(\alpha,\beta)$	$\lvert\langle\hat{A}_t\rangle_w/\lambda_-\rvert^2$
(a)	$\pi/4$	1	0
(b)	$3\pi/4 + 2.2\times10^{-2}$	-0.999	2065
(c)	$3\pi/4$	-1	indeterminate

the distribution functions (A1) and (A2) are functions of the product $g\lambda_-$, we conclude that the testing function (13) is also the UMPU test for the current problem.

The testing power of the measurement without postselection $b_{\rm nps}(\lambda_-)$ is given by Eq. (A3) and the one of the weak measurement $b_{\rm ps}(\lambda_-)$ is Eq. (A4). From these powers we obtain the relation (A5), which gives the inequality

$$b_{\rm ps}(\lambda_-) \geqslant b_{\rm nps}(\lambda_-), \tag{16}$$

which holds under the condition that the pair of α and β satisfies

$$C(\alpha,\beta) := \sin 2\alpha \sin 2\beta \leqslant 0. \tag{17}$$

Here we have used the inequality (A6). As stated in Sec. II, the amplification effect is induced if this condition (17) is satisfied. We also remark that condition (17) is related to the requirement for a weak value derived in Ref. [17] as $|\langle\hat{A}_t\rangle_w|^2 \geqslant |\lambda_-|^2 \Leftrightarrow C(\alpha,\beta) \leqslant 0$.

The experimental results are summarized in Fig. 2 in Ref. [6]. Table I shows the value of $C(\alpha,\beta)$ and the weak value $|\langle\hat{A}_t\rangle_w/\lambda_-|^2$ with the values of α and β that were tuned in the actual experiment in the three cases. We focus on two cases: (a) the measurement *de facto* without postselection[1] and (c) the weak measurement with the postselected state orthogonal to the preselected state. Case (a) exhibits the single Gaussian distribution, so we cannot distinguish whether the crystal is birefringent or not. However, in case (c), which shows the two-peak distribution, we can clearly recognize that the crystal is birefringent. We note that because the weak measurement case (b) meets the approximation $g|\langle\hat{A}\rangle_w| \ll 1$, the final probe distribution is virtually a single Gaussian distribution shifted from the initial state as considered in Ref. [1]. Case (c) gives significant amplification, which does not satisfy the approximation [4]. In both cases, the inequality (16) holds for the condition (17). Actually, case (c) gives a remarkable amplification effect as shown in Fig. 2, whereas the plots of $b_{\rm ps}$ and $b_{\rm nps}$ are almost overlapped in case (b). Thus we have shown that the angles of the two polarizers are the only factor that determines the case when the weak measurement is superior to the ordinary measurement without postselection in terms of the testing power.

IV. SUMMARY AND DISCUSSION

We have shown that the weak measurement, i.e., the measurement with the postselection, can be more powerful

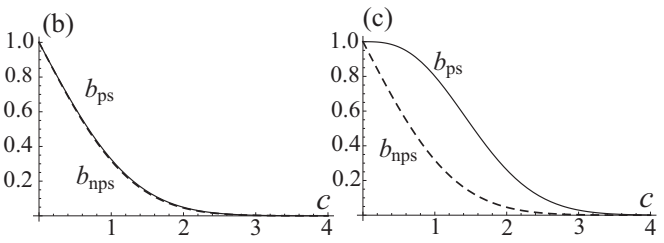

FIG. 2. Powers $b_{\rm ps}$ (solid line) and $b_{\rm nps}$ (dashed line) plotted with the critical points c as the horizontal axis in the two cases: (b) $\beta = 3\pi/4 + 2.2\times10^{-2}$ and (c) $\beta = 3\pi/4$. The other parameters are fixed as $\alpha = \pi/4$, $w_0 = 55\ \mu$m, and $g\lambda_- = 0.32\ \mu$m.

than the measurement without postselection in the hypothesis testing to determine whether or not the crystal is birefringent. When the (total) weak value given by the angles of the two polarizers is larger than the eigenvalues of the (total) observable, WVA has the advantage for the present problem. In particular, the pair of angles α and β, which does not satisfy the approximation $g|\langle\hat{A}\rangle_w| \ll 1$, gives the really powerful testing. According to the authors of Ref. [4], the amplification effect is rather striking when the approximation breaks down. Our conclusion obtained through statistical analysis supports their view on WVA. Here we have essentially treated the testing problem for the eigenvalue ($H_0 : \lambda_- = 0$ and $H_1 : \lambda_- \neq 0$), not the interaction strength ($H_0 : g = 0$ and $H_1 : g \neq 0$) that was treated in the previous work [17]. In either case, the decision function (13) gives the UMPU test and works well.

It is often argued that postselection reduces the number of detectable data, which is a statistical disadvantage of WVA [11–15]. On the other hand, the hypothesis testing generally works even if the number of detected data is small [17]. We emphasize that the experiment by Ritchie *et al.* [6] has actually shown the birefringence of the crystal by the postselection, although the detected intensity is much smaller ($\sim10^{-5}$) than that of the ordinary measurement case (Fig. 2 in Ref. [6]). We note that we observe no data with a completely orthogonal pair of α and β when the null hypothesis is really true. Then, practically, it is important to keep α and β almost orthogonal but not quite, while the approximation is not still satisfied.

In the current task we have studied the classic experiment [6], regarding it as testing the birefringence of the crystal. The experiment is a helpful example to consider the hypothesis testing with WVA [17] because it is investigated outside the validity of approximation, especially the case of postselection completely orthogonal to preselection. The hypothesis testing method can be applied to other WVA experiments, for instance, the detection of the spin Hall effect of light [7] and sensing the tilted mirror in the interferometer [8]. To clearly show the effectiveness of WVA in the experiments quoted above, we need the data for the region where the approximation ($g|\langle\hat{A}\rangle_w| \ll 1$) breaks down and the data of the ordinary measurements to compare. With those data, it would be interesting to see the effectiveness of WVA by applying our hypothesis testing method. We remark that to apply the proposed hypothesis testing method as it is, several

[1]The second polarizer slightly changes only the polarizations of the unoverlapped region, but not the one of the overlapped region.

assumptions (considered in Ref. [17]) are needed: that the probe state is given by the Gaussian profile, the measured system is described as the two-quantum-state system, and the experiment is based on the measurement of the position or the real part of the weak value. If we want to test by measuring the momentum or the imaginary part of the weak value, we should establish the appropriate testing function for the experiment.

ACKNOWLEDGMENT

The author thanks Professor A. Hosoya for valuable comments and reading the manuscript.

APPENDIX

Here we provide some equations to derive the inequality (16). From the distributions (7) and (10) we can obtain the adjusted distributions as

$$f_{\rm nps}^{\rm adj}(y|\lambda_-) = f_{\rm nps}(y+g\lambda_+|\lambda_H,\lambda_V) = \frac{1}{\sqrt{2\pi w_0^2}}\left(\cos^2\alpha e^{-(y-g\lambda_-)^2/2w_0^2} + \sin^2\alpha e^{-(y+g\lambda_-)^2/2w_0^2}\right), \tag{A1}$$

$$\begin{aligned} f_{\rm ps}^{\rm adj}(y|\lambda_-) &= f_{\rm ps}(y+g\lambda_+|\lambda_H,\lambda_V) \\ &= \frac{\cos^2\alpha\cos^2\beta\, e^{-(y-g\lambda_-)^2/2w_0^2} + \sin^2\alpha\sin^2\beta\, e^{-(y+g\lambda_-)^2/2w_0^2} + (1/2)\sin 2\alpha\sin 2\beta e^{-(y^2+g^2\lambda_-^2)/2w_0^2}}{\sqrt{2\pi w_0^2}\left(\cos^2\alpha\cos^2\beta + \sin^2\alpha\sin^2\beta + (1/2)\sin 2\alpha\sin 2\beta\, e^{-g^2\lambda_-^2/2w_0^2}\right)}. \end{aligned} \tag{A2}$$

On the basis of the decision function (13) and the adjusted distributions (A1) and (A2), we can calculate the testing power (12) of the measurement without postselection as

$$b_{\rm nps}(\lambda_-) = 1 - \frac{1}{2}\left(\mathrm{erf}\left[\frac{cw_0 - g\lambda_-}{\sqrt{2w_0^2}}\right] + \mathrm{erf}\left[\frac{cw_0 + g\lambda_-}{\sqrt{2w_0^2}}\right]\right) \tag{A3}$$

and the one of the weak measurement as

$$b_{\rm ps}(\lambda_-) = 1 - \frac{(\cos^2\alpha\cos^2\beta + \sin^2\alpha\sin^2\beta)\left(\mathrm{erf}\left[\frac{cw_0-g\lambda_-}{\sqrt{2w_0^2}}\right] + \mathrm{erf}\left[\frac{cw_0+g\lambda_-}{\sqrt{2w_0^2}}\right]\right) + \sin 2\alpha\sin 2\beta\, e^{-g^2\lambda_-^2/2w_0^2}\mathrm{erf}\left[\frac{c}{\sqrt{2}}\right]}{2(\cos^2\alpha\cos^2\beta + \sin^2\alpha\sin^2\beta) + \sin 2\alpha\sin 2\beta\, e^{-g^2\lambda_-^2/2w_0^2}}. \tag{A4}$$

From these we obtain the relation

$$\frac{1-b_{\rm ps}(\lambda_-)}{1-b_{\rm nps}(\lambda_-)} - 1 = \frac{\sin 2\alpha\sin 2\beta\, e^{-g^2\lambda_-^2/2w_0^2}\left(\frac{2\mathrm{erf}[c/\sqrt{2}]}{\mathrm{erf}[(cw_0-g\lambda_-)/\sqrt{2w_0^2}]+\mathrm{erf}[(cw_0+g\lambda_-)/\sqrt{2w_0^2}]} - 1\right)}{2(\cos^2\alpha\cos^2\beta + \sin^2\alpha\sin^2\beta) + \sin 2\alpha\sin 2\beta\, e^{-g^2\lambda_-^2/2w_0^2}}. \tag{A5}$$

We can derive the inequality (16) for Eq. (17) by using the inequality

$$\frac{2\mathrm{erf}[c/\sqrt{2}]}{\mathrm{erf}\left[(cw_0 - g\lambda_-)/\sqrt{2w_0^2}\right] + \mathrm{erf}\left[(cw_0 + g\lambda_-)/\sqrt{2w_0^2}\right]} - 1 > 0, \tag{A6}$$

which holds for $g\lambda_- \neq 0$ and is shown in Appendix A of Ref. [17].

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, How the Result of a Measurement of a Component of the Spin of a Spin-1/2 Particle Can Turn Out to be 100, Phys. Rev. Lett. **60**, 1351 (1988).

[2] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Colloquium: Understanding quantum weak values: Basics and applications, Rev. Mod. Phys. **86**, 307 (2014).

[3] A. M. Steinberg, Quantum measurement: A light touch, Nature (London) **463**, 890 (2010).

[4] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, The sense in which a "weak measurement" of a spin-1/2 particle's spin component yields a value 100, Phys. Rev. D **40**, 2112 (1989).

[5] R. Jozsa, Complex weak values in quantum measurement, Phys. Rev. A **76**, 044103 (2007).

[6] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Realization of a Measurement of a "Weak Value", Phys. Rev. Lett. **66**, 1107 (1991).

[7] O. Hosten and P. Kwiat, Observation of the spin Hall effect of light via weak measurements, Science **319**, 787 (2008).

[8] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Ultrasensitive Beam Deflection Measurement via Interferometric Weak Value Amplification, Phys. Rev. Lett. **102**, 173601 (2009).

[9] A. N. Jordan, J. Martinez-Rincon, and J. C. Howell, Technical Advantages for Weak-Value Amplification: When Less Is More, Phys. Rev. X **4**, 011031 (2014).

[10] J. Lee and I. Tsutsui, Merit of amplification by weak measurement in view of measurement uncertainty, Quantum Stud.: Math. Found. **1**, 65 (2014).

[11] G. C. Knee, G. A. D. Briggs, S. C. Benjamin, and E. M. Gauger, Quantum sensors based on weak-value amplification cannot overcome decoherence, Phys. Rev. A **87**, 012115 (2013).

[12] S. Tanaka and N. Yamamoto, Information amplification via postselection: A parameter-estimation perspective, Phys. Rev. A **88**, 042116 (2013).

[13] C. Ferrie and J. Combes, Weak Value Amplification is Suboptimal for Estimation and Detection, Phys. Rev. Lett. **112**, 040406 (2014).

[14] G. C. Knee and E. M. Gauger, When Amplification with Weak Values Fails to Suppress Technical Noise, Phys. Rev. X **4**, 011032 (2014).

[15] G. C. Knee, J. Combes, C. Ferrie, and E. M. Gauger, Weak-value amplification: State of play, arXiv:1410.6252.

[16] L. Vaidman, Comment on "Weak value amplification is suboptimal for estimation and detection", arXiv:1402.0199.

[17] Y. Susa and S. Tanaka, Statistical hypothesis testing by weak-value amplification: Proposal and evaluation, Phys. Rev. A **92**, 012112 (2015).

[18] T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach* (Academic, New York, 1967).