Statistical Hypothesis Testing by Weak-Value Amplification: Proposal and Evaluation

Yuki Susa\textsuperscript{1} and Saki Tanaka\textsuperscript{2}

\textsuperscript{1}Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
\textsuperscript{2}Graduate School of Science and Technology, Keio University, Yokohama, Kanagawa, Japan

(Dated: February 24, 2015)

We study the detection capability of the weak-value amplification on the basis of the statistical hypothesis testing. We propose a reasonable testing method in the physical and statistical senses to find that the weak measurement with the large weak value has the advantage to reduce the possibility of missing the presence of interaction. Our present work overcomes the critical dilemma of the weak-value amplification that the larger the amplification is, the smaller the number of data becomes, because the statistical hypothesis testing works even for a small number of data. This is contrasted with the parameter estimation by the weak-value amplification in the literature which requires a large number of data.

PACS numbers: 03.65.Ta, 03.67.-a

I. INTRODUCTION

The “weak-value amplification” has been studied as a promising technique for improving an accuracy of a precision measurement \cite{1,2}. The concept of the weak value comes from the weak measurement which was proposed by Aharonov and his coworkers in 1988 \cite{3–5}. Originally, the weak measurement was introduced as an example of the two-state-vector formalism for intuitive understanding of the time irreversibility of a measurement in a quantum system \cite{6}, where the “weak” means the weak coupling between the two quantum systems, the measured system and the measuring probe. Usually, we assume that the interaction Hamiltonian is of the von Neumann type which gives a displacement of the order of the coupling constant to the probe distribution \cite{7}. An important point of the weak measurement is the postselection of the measured system state after the interaction. By this operation, the weak value shows up as the shift of the expectation value of the probe position or momentum. The weak value can be outside the eigenvalue range of the system observable by choosing the postselected state of the system almost orthogonal to its initial state \cite{8}. We call this effect the weak-value amplification (WVA). If we choose an appropriate postselected state, the shift of the expectation value of probe position or momentum enhanced larger than the shift given in the measurement without postselection which coincides with the coupling constant. Therefore, we may hope to extract the information about the coupling constant even if the constant is smaller than a noise.

Reference \cite{9} introduced the basic concept of the weak value and its application. It is known that the amplified shift has an upper bound if we take Gaussian for the initial probe state \cite{10}. Especially, when the measured system is a two-state system, some researchers have shown the upper bound analytically without any approximation \cite{11,13}. The initial measuring probe wave function which maximizes the amplification factor was studied in Refs. \cite{14,16}. The WVA was confirmed in several experimental studies. Setting two polarizers in front and behind of a birefringent crystal, the weak value was measured and found to become large by arranging two polarizers almost orthogonal \cite{17}. Hosten and Kwiat observed the spin Hall effect of light with the WVA \cite{18}. Dixon et al. monitored the laser deflection by the small tilting Piezo mirror in the Sagnac interferometer and measured the angle of the mirror in precision \cite{19}. Viza et al. demonstrated the velocity measurement of the longitudinal moving mirror in the Michelson interferometer \cite{20}.

There are theoretical researches about technical utilities of the WVA. Nishizawa et al. compared the signal and the shot-noise in an optical interferometer \cite{21}. Jordan et al. showed that the WVA has the error tolerance of the systematic error \cite{22}. Lee and Tsutsui discussed the causes of the errors in the weak measurement with finite data and the merit of the WVA \cite{23}. Knee and Gauger showed that the WVA is helpful when the resolution power of a detector is low \cite{24}. However, the problem is that the larger the amplification factor is, the smaller the success probability of the postselection becomes leading to the possible disadvantage of the WVA as some papers argue \cite{24,25}. The failed postselection makes data loss and it reduces the Fisher information which is an important factor to estimation accuracy, since the inverse of the Fisher information gives the lower bound of the mean squared error of the accuracy, and these are equal for the infinite number of data \cite{28}. Because a large number of data is necessary for an estimator to approach the true value, the weak measurement is not suitable for the coupling constant estimation, when the number of experiments is limited. Reference \cite{29} gives the summary of this point. On the other hand, some researchers have mentioned that the data loss by postselection is not usually needed to consider in practical cases \cite{11,30}.

Lately the interaction detection capability of the WVA has been focused on. In this problem, using the WVA, we
want to decide whether the interaction exists or not in an indirect quantum measurement process. To study such a problem in statistics, the hypothesis testing is a well-known method. The hypothesis testing does not require an infinite number of data like the estimation and the accuracy of testing is normally independent of the number of data. Therefore, we note that the hypothesis testing can be used for a small number of data given by the weak measurement. In Ref. [27], the authors claim that the WVA is suboptimal for the interaction detection. Their conclusion is based on the discussion of the likelihood-ratio test with some debate [30–32]. It seems to us, however, we have to pay more attention to the well-known fact that the likelihood-ratio test is not appropriate to the detection of the presence of the interaction. The Neyman-Pearson lemma shows that the likelihood-ratio test becomes a uniformly most powerful (UMP) test, if the hypotheses are both simple [33]. However, the interaction existence testing needs the discussion with a composite hypothesis. At least, the likelihood-ratio test is not optimal test method for the interaction existence detection, it is generally secondary. More precisely, because the interaction detection problem is a two-sided test, it is not a proper attempt to obtain the UMP test. It is often possible to find a better test in a very specific case. For a two-sided test, a uniformly most powerful unbiased (UMPU) test is the standard procedure.

In this paper, we propose a statistical hypothesis testing to replace the likelihood-ratio test and analytically evaluate the interaction detection capability of the WVA. To determine whether the interaction exists or not, we set a decision function such as the measurement outcome divided by the initial fluctuation of the measuring probe distribution (|x|/σ below Eq. (16)). Under this decision function and the particular condition for the weak value, we find that the WVA can supersede the ordinary measurement without postselection, which does not depend on the significance level, the critical point nor the interaction strength. We emphasize that the advantage of the WVA is the reducing the possibility of missing the presence of the interaction with the false alarm rate fixed. Throughout this paper, we assume that the initial probe wave function is Gaussian and the measured system is a two-state system as considered in Refs. [11][12].

Our paper is organized as follows. In Sec. [II] we give brief reviews of the weak measurement and the statistical hypothesis testing which includes a UMPU test for the two-sided test. Section [III] gives the main result of this work. We propose a proper decision function for the interaction detection and evaluate the statistical errors so-called “type 1 error” and “type 2 error”. We show the superiority of the WVA and the required condition for the weak value with analytical derivation. In Sec. [IV] we consider an additive white Gaussian noise model. In this situation, the result of Sec. [III] holds when the unknown extra fluctuation exists. We have summary and discussion in Sec. [V]. There, we discuss the case that the data would be unobtainable by failure of the postselection. Some complicated calculations and the supplementary discussion of the unobtainable case are shown in Appendixes. We use the unit h = 1 and the subscripts that “ps” means the case of the measurement with the postselection i.e., the weak measurement and “nps” indicates the case of the ordinary measurement i.e., the measurement with no postselection.

II. REVIEW OF WEAK MEASUREMENT AND HYPOTHESIS TESTING

A. Weak measurement and probability distributions

We recapitulate the standard weak measurement process [2][3][7][9] and derive the probability distribution after the measurement. The weak measurement is described as the indirect quantum measurement formalism. To carry out the weak measurement, initially we prepare a measured system H and a measuring probe K. (Hereafter we omit the indexes H and K unless otherwise stated.) The initial state of the measured system H is a preselected state |i⟩_H and the initial state of the probe system K is |ψ⟩ = ∫ ψ(x)|x⟩dx, which is taken as the Gaussian profile:

$$\psi(x) = Ce^{-\frac{x^2}{2\sigma^2}}, \quad C^2 = \frac{1}{\sqrt{2\pi\sigma^2}}. \quad (1)$$

where the x represents the position of the probe. We denote $\hat{\rho}_i := |\psi⟩⟨\psi|$ for the probe state in a density matrix expression for later convenience. We assume the von Neumann interaction so that the time evolution operator given by $\hat{U} = \exp(-igA^H \otimes \hat{p}^K)$, where $A^H$ is an observable defined in H and $\hat{p}^K$ is a momentum operator satisfying $[x^K, \hat{p}^K] = i$ defined in K. The parameter g indicates an unknown coupling strength. The interaction produces the state of the combined system as

$$\hat{\rho}_{\text{int}}^{H \otimes K} = \hat{U}(|i⟩_H⟨i| \otimes \hat{p}^K)\hat{U}^\dagger. \quad (2)$$

Finally, we postselect the measured system state $|f⟩_H$.

Here we assume that the interaction strength g is sufficiently small that the first approximation in g is available. The postselection translates the wave function by $g\text{Re}(\hat{A})_w$ i.e.,

$$\psi(x) \rightarrow \psi(x - g\text{Re}(\hat{A})_w), \quad (3)$$

where the $\langle \hat{A} \rangle_w$ is called the weak value defined by

$$\langle \hat{A} \rangle_w := \langle f|\hat{A}|i⟩ \langle f|i⟩. \quad (4)$$

We can easily see that the weak value has a generally complex value, and becomes infinitely large when the postselected state $|f⟩$ is almost orthogonal to the initial state $|i⟩$. Because of the large $\langle \hat{A} \rangle_w$, we find that the
shift of the probe wave function gets large as can be seen from Eq. (3). This implies that the coupling constant \( g \) is effectively amplified. This effect is called “weak-value amplification” (WVA) \[18\]. We emphasize that the post-selection is essential in the WVA.

For the later discussion, we carry out full order calculation of the final probability distribution of the measuring probe concentrating on the two-state system case for the measured system. The final probe state becomes

\[
\hat{\rho}_{ps}^K := \frac{\text{Tr}_H\left((|f\rangle_H \langle f| \otimes \hat{I}^K)\hat{\rho}_{\text{int}}^{H \otimes K}\right)}{\text{Tr}\left[(|f\rangle_H \langle f| \otimes \hat{I}^K)\hat{\rho}_{\text{int}}^{H \otimes K}\right]}
\]

(5)

Here we note that the denominator coincides with the amplification “weak-value amplification” (WVA) \[18\]. We emphasize that the post-selection causes the difference from \( f_{\text{ps}} \) to \( f_{\text{ps}} \). The weak value \( \langle \hat{A} \rangle_w \) plays an important role in determining the distribution change. Note that the condition \( \text{Re}(\langle \hat{A} \rangle_w) = 0 \) makes the two peaks at the same height and that the increases of the imaginary part \( \text{Im}(\langle \hat{A} \rangle_w) \) depresses the dip deeper while increasing the distance of the two peaks. This may not be very obvious, but is justified by the closer algebraic examination of the expression (6). With an appropriate postselection, we can obtain peak shifts larger than the ordinary peak-to-peak distance \[8\], which is given by the coupling constant times the eigenvalues of the observable as indicated in Fig. 1. Note that in the case of \( \langle \hat{A} \rangle_w = \pm 1 \), the postselection reduces to the projective measurement of \( \hat{A} \). A similar discussion goes through for \( |i\rangle = |+\rangle \) or \( |\rangle \) in the case without postselection.

\[\text{FIG. 1: A sketch of the distribution functions when } \sigma = 1 \text{ and } g = 1.5. \text{ The blue curve is the probe distribution before the interaction. The red and the yellow curves are the ones after the interaction with postselection and without postselection, respectively. We choose the initial state of the measured system as } |i\rangle = (|+\rangle + |\rangle)/\sqrt{2}, \text{ and the postselected state which satisfies the weak value } \text{Re}(\langle \hat{A} \rangle_w) = 0 \text{ and } |\langle \hat{A} \rangle_w| = 5.\]

\[\text{B. Hypothesis testing for detection problem}\]

Generally, the statistical hypothesis testing is used for a mathematical decision from the data of measurement \[34, 35\]. Our aim of this paper is the evaluation of the interaction detection capability by the WVA in which the size of data is not necessarily large. In the classical theory of statistics, the hypothesis testing provides a better solution for such evaluation than the estimation theory as explained below. In the estimation theory, it is known that the mean squared error of the estimator approaches the inverse of the Fisher information at infinite number of data. Then, the Fisher information gives a good description for the accuracy of the estimation with a large number of data \[28, 37, 39\], while the hypothesis testing can be evaluated by the error probability regardless of the number of data. We emphasize that our discussion based on the hypothesis testing theory works even when we have a small number of data.

We take the hypotheses as follows: (a) the null hy-
hypothesis $H_0$: absence of the interaction (i.e., $g = 0$), (b) the alternative hypothesis $H_1$: presence of the interaction (i.e., $g \neq 0$). This hypothesis testing problem is a two-sided test in the sense that the estimated parameter deviates in either direction ($g > 0$, $g < 0$) from the null hypothesis ($g = 0$). We derive a decision function which takes the binary value as $d(x) = 0$ or 1. If it is 0, we reject the alternative hypothesis, and if it is 1, we reject the null one. A range of measurement results such that $d(x) = 0$ is called the acceptance region. When the measurement result falls within this region, the null hypothesis is rejected. Since the decision function $d(x)$ is independent of the coupling parameter $g$, sometimes we get the wrong indications from the decision function.

The such wrong indications are classified to the two types, (i) the type 1 error: even if the null hypothesis is true, we wrongly reject it, (ii) the type 2 error: even if the null hypothesis is false, we wrongly accept it as truth. In our case, the type 1 error represents “there is no interaction but we wrongly guess the interaction exists i.e., falsely alarmed”. The type 2 error means “the interaction indeed exists, but we wrongly guess the interaction does not exist i.e., miss the presence of the interaction”. The probability of the type 1 error is calculated as the integration of the probability distribution function over the rejection region of the null hypothesis when the null hypothesis is true. The probability of the type 2 error is calculated as the integration of the probability distribution function over the acceptance region of the null hypothesis when the alternative hypothesis is true. The smaller probabilities of these errors become, the better the testing is. Since it is difficult to make the two types of error small simultaneously, we make the probability of the type 2 error as small as possible while suppressing the probability of the type 1 error under a certain significance level as the general strategy [34]. In this paper, we compare the weak measurement and the measurement without postselection by evaluating the probabilities of the two types of error for each measurement.

The uniformly most powerful (UMP) test is one of the good testings in the general strategy [36]. To carry out such a test, an adequate decision function is necessary. The Neyman-Pearson lemma is a famous example for explaining the UMP test. This lemma claims the likelihood-ratio test is UMP only when the both hypotheses are simple. That is, the decision function

$$d(x) = \begin{cases} 0, & \text{if } f(x|\theta_1)/f(x|\theta_0) < c \\ r, & \text{if } f(x|\theta_1)/f(x|\theta_0) = c \\ 1, & \text{if } f(x|\theta_1)/f(x|\theta_0) > c \end{cases} \quad (9)$$

gives the UMP test when the null hypothesis is $H_0: \theta = \theta_0$ and the alternative is $H_1: \theta = \theta_1$. Here, $c$ is a real number and called a critical point. At the critical point, the $d(x)$ becomes the randomized test $d(x) = r$, and $r$ ($0 \leq r \leq 1$) is the probability to accept the

null hypothesis [40]. It is important to note that the likelihood-ratio test is not an appropriate testing when either or both of the hypotheses is composite, for instance our hypotheses. To avoid this problem, the maximum likelihood-estimator (MLE) is commonly used, e.g., see Ferrie and Combes [23]. However, from the statistical inference [28, 33], it is widely known that the MLE without the large number of data is not helpful for extracting the information of a physical system.

Generally speaking, the UMP test does not exist in the two-sided test such as $H_0: \theta = \theta_0$ and $H_1: \theta \neq \theta_0$ including our problem. To find a good test in the two-side test, the concept of the unbiased test $d(x)$ is needed satisfying

$$\beta(\theta) := \int d(x)f(x|\theta)dx \geq \alpha, \forall \theta \in \Theta_1, \quad (10)$$

where the $\alpha$ is a significance level and $\Theta_1$ is a set of parameters delineating the alternative hypothesis $H_1$. We note that from the $\beta(\theta)$ called statistical power, we can calculate the probabilities of each error as follows: one of the type 1 error is $Pr[E_1] := \beta(\theta_0)$ and one of the type 2 error is $Pr[E_2] := 1 - \beta(\theta \in \Theta_1)$. According to the following lemma, we can obtain the UMPU test which is a good test for composite hypotheses such as a two-sided test [34].

**Lemma.** If the hypotheses are given by the two-sided test: $H_0: \theta = \theta_0$ and $H_1: \theta \neq \theta_0$, we assume that the decision function $d(x)$ satisfies

$$\partial_0 \beta(\theta) = \int d(x)\partial_0 f(x|\theta)dx, \quad (11)$$
$$\beta(\theta_0) = \alpha, \quad (12)$$
$$\partial_0 \beta(\theta_0) = 0. \quad (13)$$

For an any fixed $\theta_1$ such that $\theta_1 \neq \theta_0$, if the test $d(x)$ is given by

$$d(x) = \begin{cases} 0, & \text{if } F(x) < 0 \\ r, & \text{if } F(x) = 0 \\ 1, & \text{if } F(x) > 0 \end{cases} \quad (14)$$

where

$$F(x) := f(x|\theta_1) - c_1 f(x|\theta_0) - c_2 \partial_0 f(x|\theta)|_{\theta=\theta_0} \quad (15)$$

with certain parameters $c_1$ and $c_2$, the $d(x)$ is the UMPU test.

This lemma works for a small number of samples such as the data given by the weak measurement. In the following section, we propose a test which has a physical meaning, and we check that the test is UMPU on the basis of the above lemma.
III. HYPOTHESIS TESTING WITH WEAK-VALUE AMPLIFICATION

A. Merit of WVA in interaction detection

In what follows, we derive the UMPU test for the detection of the presence of the weak interaction $g$ in the two cases. As mentioned in Sec. IIIB, our proposed test is the best one among all unbiased tests. Then we compare the best test in the WVA with that in the measurement without postselection by explicit forms. We evaluate the testing capability, comparing the probabilities for the type 1 and the type 2 errors for the two cases: the weak measurement and the measurement without postselection. We remark that, in this section, we treat the measurement result is a positive constant that we choose as we like. The rejection region is fixed to $|x| > c \sigma$. We verify that this decision function (16) is the UMPU test for such a one-sided test, and we can show our decision function (16) gives a UMP test (see Sec. IIIIB for detail).

Here, we compare the probabilities of the type 1 and 2 errors obtained by Eqs. (19) and (20). First, we consider the type 1 error when the coupling constant is $g = 0$ and the measurement result is $|x| > c \sigma$. Since the distribution functions $f_{ps}(x|g)$ and $f_{nps}(x|g)$ coincide at $g = 0$;

$$f_{ps}(x|g = 0) = f_{nps}(x|g = 0) = e^{-\frac{x^2}{2\sigma^2}/\sqrt{2\pi\sigma^2}},$$  \hspace{1cm} (17)

the probabilities of the type 1 error of $f_{ps}(x|g)$ and of $f_{nps}(x|g)$ are the same as

$$\Pr[\mathcal{E}_1] = \beta(0) = 1 - \text{erf}\left[\frac{c}{\sqrt{2}}\right],$$  \hspace{1cm} (18)

where $\text{erf}[x] := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the error function. The $\Pr[\mathcal{E}_1]$ can be any significance level by choosing $c$. Hence, this test suits the standard strategy [28].

Next, we evaluate the probabilities of the type 2 error when $g \neq 0$ and $|x| < c \sigma$. In this case, the distribution functions (19) and (20) are different. The probability of the type 2 error in the weak measurement is

$$\Pr[\mathcal{E}_{2,ps}] = 1 - \beta_{ps}(g) = \frac{1}{2} \left[ 1 + \left| \langle \hat{A} \rangle_w \right|^2 + \left( 1 - \left| \langle \hat{A} \rangle_w \right|^2 \right) e^{-\frac{c^2}{2\sigma^2}} \right].$$  \hspace{1cm} (19)

In the measurement without postselection, the probability of the type 2 error is given by

$$\Pr[\mathcal{E}_{2,nps}] = 1 - \beta_{nps}(g) = \frac{1}{2} \left[ \text{erf}\left[\frac{c \sigma - g}{\sqrt{2\sigma^2}}\right] + \text{erf}\left[\frac{c \sigma + g}{\sqrt{2\sigma^2}}\right] \right].$$  \hspace{1cm} (20)

To determine which measurement gives more benefit, it is enough to compare the probabilities (19) and (20). We can derive the following equation by arranging the ratio $\Pr[\mathcal{E}_{2,ps}] / \Pr[\mathcal{E}_{2,nps}]$, which is symmetric under the sign...
change $g \leftrightarrow -g$, as
\[
\Pr[\mathcal{E}_{2,ps}] / \Pr[\mathcal{E}_{2,nps}] - 1 = \frac{(1 - |\langle A \rangle_w|^2) \left( \frac{2e^{-\frac{x^2}{2\sigma^2}}}{\text{erf} \left( \frac{x}{\sqrt{2\sigma^2}} \right) + \text{erf} \left( \frac{-x}{\sqrt{2\sigma^2}} \right)} - 1 \right)}{1 + |\langle A \rangle_w|^2 + \left( 1 - |\langle A \rangle_w|^2 \right) e^{-\frac{g^2}{2\sigma^2}}}.
\]
(21)

We can see that the inequality
\[
\Pr[\mathcal{E}_{2,ps}] \leq \Pr[\mathcal{E}_{2,nps}]
\]
holds for such a weak value that $|\langle A \rangle_w| \geq 1$ as shown in Appendix A. Therefore, the probability of the type 2 error with the postselection can be less than the one without postselection in a certain case.

Figure 2 shows the ratio $\Pr[\mathcal{E}_{2,ps}] / \Pr[\mathcal{E}_{2,nps}]$ for the three parameters $|\langle A \rangle_w|$, $g$, and $c$. We can see from these graphs that the inequality (22) holds when the weak value $|\langle A \rangle_w|$ is larger than 1. They also indicate that when the coupling constant $g$ and the critical point $c$ are relatively small, the WVA well works to decrease the probability of the type 2 error. We believe that this is the heart of the weak-value amplification which is mathematically well-grounded. On the other hand, if the $g$ is large, the large weak value is as helpful to detect as the ordinary measurement. This property comes from the difference of the distribution functions $f_{ps}$ and $f_{nps}$. The difference can be explained in Fig. 3. When the $g$ is small as shown in Fig. 3(a), there is a big difference between $f_{ps}$ and $f_{nps}$ for the small $|x|$. Note that there is no difference between $f_{ps}$ and $f_{nps}$ for a large $g$ as we can see from Fig. 3 (b) and (c). The probability of the type 2 error is given by the integration over the interval $[-c\sigma, c\sigma]$. Since $f_{ps}$ is smaller than $f_{nps}$ in the central region of $x$ for a fixed small $g$, the ratio $\Pr[\mathcal{E}_{2,ps}] / \Pr[\mathcal{E}_{2,nps}]$ becomes small, if we properly choose the critical point $c$.

To summarize this section, if $|\langle A \rangle_w| \geq 1$, the WVA has the advantage for the reduction of the type 2 error while keeping the type 1 error fixed, regardless of the coupling constant $g$, the initial fluctuation $\sigma$ and the critical point $c$. Namely, the weak measurement more reduces the possibility of missing the presence of the interaction with the false alarm rate fixed than the ordinary measurement i.e., the measurement without postselection, if the weak value is outside of the normal range of the eigenvalues.

B. A proof that our test is UMPU

We prove that our decision function (16) is the UMPU test for the probability distributions given by the each measurement in accordance with Lemma giving the test (13) and (15). We can easily see that the functions (6) and (8) satisfy Eq. (11). We have shown Eq. (18) which indicates that Eq. (12) can be satisfied by appropriately choosing the critical point $c$ in each measurement. We can also show that Eq. (13) is satisfied as
\[
\partial_g \beta_{ps}(g) \big|_{g=0} = \partial_g \beta_{nps}(g) \big|_{g=0} = \frac{1}{\sqrt{2\pi\sigma^2}} \left( e^{\frac{(c+g)^2}{2\sigma^2}} - e^{\frac{(c-g)^2}{2\sigma^2}} \right) \bigg\vert_{g=0} = 0.
\]
(23)

and
\[
\partial_g \beta_{nps}(g) \big|_{g=0} = \partial_g (1 - \Pr[\mathcal{E}_{2,nps}]) \big|_{g=0} = \frac{1}{\sqrt{2\pi\sigma^2}} \left( e^{\frac{(c+g)^2}{2\sigma^2}} - e^{\frac{(c-g)^2}{2\sigma^2}} \right) \bigg\vert_{g=0} = 0.
\]
(24)

Here, we consider the case of the weak measurement. Equation (15) with $f_{ps}(x|g)$ becomes
\[
\frac{F_{ps}(x)}{f_{ps}(x|0)} = \mathcal{G}_{ps}(x) - c_{1,ps} - c_{2,ps} \frac{\text{Re}(\hat{A})_w}{\sigma^2} x,
\]
(25)
where

\[
G_{nps}(x) := \frac{e^{-\frac{x^2}{2\sigma^2}}}{2 \left[ 1 + |\langle \hat{A} \rangle_w|^2 + (1 - |\langle \hat{A} \rangle_w|^2) e^{-\frac{x^2}{2\sigma^2}} \right]^2} \left\{ \left( 1 + |\langle \hat{A} \rangle_w|^2 + 2\text{Re}(\langle \hat{A} \rangle_w) \right) e^{\frac{x^2}{2\sigma^2}} \right. \\
\left. + \left( 1 + |\langle \hat{A} \rangle_w|^2 - 2\text{Re}(\langle \hat{A} \rangle_w) \right) e^{-\frac{x^2}{2\sigma^2}} \right\} + 2 \left( 1 - |\langle \hat{A} \rangle_w|^2 \right) e^{-\frac{x^2}{2\sigma^2}}.
\]

Then we can find \(c_{1,\text{p}}} and \(c_{2,\text{p}}\) which give \(x = \pm \sigma \) for the solutions of \(F_{nps}(x)/f_{nps}(x(0)) = 0\) as below;

\[
c_{1,\text{p}} = \frac{G_{nps}(\sigma) + G_{nps}(-\sigma)}{2} = \frac{e^{-\frac{\sigma^2}{2\sigma^2}}}{1 + |\langle \hat{A} \rangle_w|^2 + (1 - |\langle \hat{A} \rangle_w|^2) e^{-\frac{x^2}{2\sigma^2}}} \left[ 2 \left( 1 + |\langle \hat{A} \rangle_w|^2 \right) \text{cosh} \left( \frac{\sigma^2}{\sigma} \right) \right], \quad (27)
\]

\[
c_{2,\text{p}} = \frac{\sigma G_{nps}(\sigma) - G_{nps}(-\sigma)}{2 \text{Re}(\langle \hat{A} \rangle_w)c} = \frac{2\sigma e^{-\frac{\sigma^2}{2\sigma^2}} \sinh \left( \frac{\sigma^2}{\sigma} \right)}{c \left[ 1 + |\langle \hat{A} \rangle_w|^2 + (1 - |\langle \hat{A} \rangle_w|^2) e^{-\frac{x^2}{2\sigma^2}} \right]}, \quad (28)
\]

for \(\text{Re}(\langle \hat{A} \rangle_w) \neq 0\). Because the \(G_{nps}(x)\) is a convex function, we can interpret the UMP test given by Lemma as the proposed decision function \(16\). Therefore, the test \(16\) with \(f_{nps}(x|g)\) is UMP. Even if \(\text{Re}(\langle \hat{A} \rangle_w) = 0\), the discussion remains valid.

A similar discussion holds in the case of the measurement without postselection. In this case, the distribution function is \(f_{nps}(x|g)\) and Eq. \(16\) becomes

\[
\frac{F_{nps}(x)}{f_{nps}(x(0))} = G_{nps}(x) - c_{1,nps} - c_{2,nps} \frac{|\langle +|i \rangle^2 - |\langle -|i \rangle^2}{\sigma^2},
\]

where

\[
G_{nps}(x) := e^{-\frac{x^2}{2\sigma^2}} \left( |\langle +|i \rangle^2 e^{\frac{\sigma^2}{\sigma}} + |\langle -|i \rangle^2 e^{-\frac{\sigma^2}{\sigma}} \right). \quad (30)
\]

Then we can obtain \(c_{1,nps}\) and \(c_{2,nps}\) which give \(x = \pm \sigma \) as the solutions of \(F_{nps}(x)/f_{nps}(x(0)) = 0\) as below;

\[
c_{1,nps} = \frac{G_{nps}(\sigma) + G_{nps}(-\sigma)}{2} = e^{-\frac{\sigma^2}{2\sigma^2}} \text{cosh} \left( \frac{\sigma^2}{\sigma} \right), \quad (31)
\]

\[
c_{2,nps} = \frac{\sigma G_{nps}(\sigma) - G_{nps}(-\sigma)}{2 \left( |\langle +|i \rangle^2 - |\langle -|i \rangle^2 \right) c} = \frac{\sigma e^{-\frac{\sigma^2}{2\sigma^2}} \sinh \left( \frac{\sigma^2}{\sigma} \right)}{c} \cdot (32)
\]

Here we have taken \(|\langle +|i \rangle|^{2} \neq |\langle -|i \rangle|^{2}\) to obtain the \(c_{2,nps}\). Even if \(|\langle +|i \rangle|^{2} = |\langle -|i \rangle|^{2}\), the discussion goes through. The \(G_{nps}(x)\) is also a convex function. Therefore, we have found that the decision function \(16\) is the UMP test in the both measurement cases.

**C. Obtaining UMP test for detection problem under certain conditions**

We show that our decision function \(16\) gives a UMP test in the case \(\text{Re}(\langle \hat{A} \rangle_w) = 0\) for the weak measurement and the case \(|\langle +|i \rangle|^{2} = |\langle -|i \rangle|^{2}\) for the measurement without postselection. In these cases, the distribution functions \(6\) and \(8\) become even, and the sign of \(g\) becomes indistinct. Then, we can practically assume the sign of \(g\) is positive without losing generality. According to the following theorem, we can obtain the UMP test for composite hypotheses such as \(H_{0} : \theta = \theta_{0}\) and \(H_{1} : \theta > \theta_{0}\).

**Theorem.** If the likelihood-ratio becomes a monotonically increasing function of the statistics \(T(x)\) which is composed of the sample data \(x\), the following test becomes the UMP for hypotheses: \(H_{0} : \theta = \theta_{0}\) and \(H_{1} : \theta > \theta_{0}\).

\[
d(x) = \begin{cases} 0 & \text{if } T(x) < c \ \\
\frac{1}{r} & \text{if } T(x) = c \\
1 & \text{if } T(x) > c \end{cases}
\]

(33)

Because Theorem is applicable to a one-sided test such as \(H_{0} : g = 0\) and \(H_{1} : g > 0\), we can show that the test \(16\) becomes UMP for the each measurement under the certain conditions which make the distribution functions after the measurement even.

We can calculate the likely- hood ratios from Eqs. \(6\) and \(8\) in the each case as

\[
\frac{f_{nps}(x|g)}{f_{nps}(x|g = 0)} = \frac{1 - \text{Im}(\langle \hat{A} \rangle_{w}^\dagger) + \left(1 + \text{Im}(\langle \hat{A} \rangle_{w}^\dagger) \right) \text{cosh} \left( \frac{\sigma^2}{\sigma} \right)}{1 - \text{Im}(\langle \hat{A} \rangle_{w}^\dagger) + \left(1 + \text{Im}(\langle \hat{A} \rangle_{w}^\dagger) \right) e^{\frac{\sigma^2}{2\sigma^2}}}, \quad (34)
\]

\[
\frac{f_{nps}(x|g)}{f_{nps}(x|g = 0)} = e^{-\frac{\sigma^2}{2\sigma^2}} \text{cosh} \left( \frac{xg}{\sigma} \right), \quad (35)
\]

respectively. To obtain them, we have used \(\text{Re}(\langle \hat{A} \rangle_w) = 0\) and \(|\langle +|i \rangle|^{2} = |\langle -|i \rangle|^{2}\). Since cosh is an even function,

\[
\text{cosh} \left( \frac{\sigma^2}{\sigma} \right) = \frac{g \cdot |x|}{\sqrt{\sigma^2}}.
\]

(36)

For the ratio \(34\) and \(35\) to be a UMP test, we demand Eqs. \(7\) and \(9\) to be functions of statistics \(T(x)\) which are independent of the unknown parameter \(g\). Then, we can set the statistics \(T(x) = |x|/\sigma\) not to contain the unknown parameter \(g\) so that the likely- hood ratios \(34\) and \(35\) become monotonically increasing functions. Thus, we find the statistics \(T(x) = |x|/\sigma\) and Theorem that our test \(16\) is UMP in the case \(\text{Re}(\langle \hat{A} \rangle_w) = 0\) for the weak measurement and the case \(|\langle +|i \rangle|^{2} = |\langle -|i \rangle|^{2}\) for the measurement without postselection.
IV. TESTING IN THE CASE WITH AN ADDITIVE WHITE GAUSSIAN NOISE

In this section, we remark on the noise tolerance of the hypothesis test proposed in the previous section. The probability distribution of the experimental result \( x \) is ideally given by the distribution functions \( \pi_0 \) or \( \pi_1 \). However, there are always noise. We assume that the noise \( y \) is added to \( x \) by passing through the device circuit with the Gaussian probability \( e^{-y^2/2s^2}/\sqrt{2\pi s^2} \) with an arbitrary fluctuation \( s \). This noise model is widely known as an additive white Gaussian noise. This is seen in the thermal noise generated in an electrical conductor or the shot noise in an electronic circuit. This Gaussian jitter noise in the estimation accuracy of the WVA was discussed in Ref. [20]. We have the distribution of \( z = x + y \) from the moment-generating functions of \( x \) and \( y \) distributions, and we show its derivation in Appendix B.

From the distribution of \( z \) we have the probabilities of the two types of errors. The probability of the type 1 error with the postselection and the one without postselection are the same as,

\[
Pr[\mathcal{E}_{1,ps}] = Pr[\mathcal{E}_{1,nps}] = 1 - \text{erf} \left( \frac{c\sigma}{2\sqrt{\sigma^2 + s^2}} \right). \tag{37}
\]

The probability of the type 2 error with the postselection is

\[
Pr[\mathcal{E}_{2,ps}] = \frac{1}{2} \left[ 1 + \frac{1}{\left( \langle A \rangle_w \right)^2 + \left( 1 - \langle A \rangle_w \right)^2 e^{-\frac{2g^2}{2\sigma^2}} \right] \cdot \left[ \left( 1 + \left( \langle A \rangle_w \right)^2 \right) \left( \text{erf} \left( \frac{c\sigma - g}{2\sqrt{\sigma^2 + s^2}} \right) + \text{erf} \left( \frac{c\sigma + g}{2\sqrt{\sigma^2 + s^2}} \right) \right) + 2 \left( 1 - \left( \langle A \rangle_w \right)^2 \right) e^{-\frac{2g^2}{2\sigma^2}} \left( \text{erf} \left( \frac{c\sigma - g}{2\sqrt{\sigma^2 + s^2}} \right) \right) \right], \tag{38}
\]

and the one without postselection is

\[
Pr[\mathcal{E}_{2,nps}] = \frac{1}{2} \left( \text{erf} \left( \frac{c\sigma - g}{2\sqrt{\sigma^2 + s^2}} \right) + \text{erf} \left( \frac{c\sigma + g}{2\sqrt{\sigma^2 + s^2}} \right) \right). \tag{39}
\]

These probabilities are almost the same as Eqs. (19) and (20), respectively except the denominator of the argument of the error functions. Comparing these probabilities as in Sec. IIIA we get the same conclusion that the WVA makes the probability of the type 2 error small when \( \langle A \rangle_w \geq 1 \). Therefore, we conclude that our testing is robust against the unknown fluctuation.

V. SUMMARY AND DISCUSSION

In this paper, we have studied the capability of the WVA to detect whether the interaction is present or not in an indirect quantum measurement scheme with the statistical hypothesis testing. We have assumed that the measured system is the two-state system and that the initial wave function of the measuring probe is Gaussian. Since the likely-hood ratio test with the MLE is not an appropriate method for the interaction detection, which should be treated as the two-sided test, we have proposed the decision function which gives the UMP test for the interaction detection. The proposal is to infer that there will be an interaction if the data distribution obtained by the measurement is outside the fluctuation of the probe distribution before the measurement. We remark that the proposed test is regarded as a UMP test in the case such that the detection problem essentially behaves as a one-sided test, because the distribution function after the measurement becomes even. The merit of the WVA is the reduction of the possibility to miss the presence of the interaction with a fixed false alarm rate than the ordinary measurement, when the absolute value of the weak value is greater than 1. We have noted that our hypothesis testing has the robustness against the additive white Gaussian noise.

We again emphasize that the number of data does not matter and our discussion holds even for a small number of measurement results. We note that our result does not conflict with the ones of the estimation theory in Refs. [22] because the estimation theory is used for accurate determination of the parameter by a large number of data.

At this stage, we need to discuss the case that we cannot obtain any measurement data due to complete failure of the postselection. To cope with such a case, we consider a makeshift decision function as an attempt to discuss in Appendix C. There, we have found that the optimal condition for reducing the type 2 error with the type 1 error under a certain significance level is that the postselected state is the eigenstate of the measured observable and the postselection is not necessary. However, this discussion has defects on the treatment of the failure of the postselection. There is no reasonable ground that we regard the failure of the postselection as the absence of interaction, because the postselection can fail whether the interaction is present or absent. When there is a case that we cannot obtain any data, the problem of the null result arises. Even though we know how to treat the null result in the projective measurement [44, 45], that of the null result in the weak measurement has not been developed yet. This remains an open problem.

Generally speaking, a UMP test and a UMPU test do not always provide an optimal solution and it is difficult to optimize the statistical hypothesis testing [44].

Acknowledgment

We appreciate for valuable comments from statistical viewpoint by Professor Fuyuhiko Tanaka. We thank Professor Akio Hosoya for reading of manuscript. Y. S. is supported by JSPS (Grant No. 25008633).

Appendix A: Proof of the inequality \( (22) \)

Here, we prove that the inequality \( (22) \) as \( Pr[\mathcal{E}_{2,ps}] \leq Pr[\mathcal{E}_{2,nps}] \) when \( \langle A \rangle_w \geq 1 \) by looking at the right hand side of Eq. (21). More precisely, we show

\[
\frac{2\text{erf} \left( \frac{c\sigma - g}{\sqrt{2\sigma^2}} \right)}{\text{erf} \left( \frac{c\sigma - g}{\sqrt{2\sigma^2}} \right) + \text{erf} \left( \frac{c\sigma + g}{\sqrt{2\sigma^2}} \right)} > 1. \tag{A1}
\]

It is enough to show \( (A1) \) only for the case \( g > 0 \), because the symmetry of the left hand side of the inequality \( (A1) \) under the exchange \( g \leftrightarrow -g \). For the case \( 0 < g \leq c\sigma \),
\[ e^{-(t-g/\sqrt{2\sigma})^2} > e^{-t^2} \] holds when \( t \geq c/\sqrt{2} \). Then we have
\[
\int_{-\infty}^{\infty} e^{-\left(\frac{t-g}{\sqrt{2\sigma}}\right)^2} dt > \int_{-\infty}^{\infty} e^{-t^2} dt
\]
\[
\Rightarrow \int_{-\infty}^{\infty} e^{-t^2} dt > \int_{-\infty}^{\infty} e^{-\left(\frac{t-g}{\sqrt{2\sigma}}\right)^2} dt
\]
\[
\Rightarrow \text{erf} \left[ \frac{c}{\sqrt{2}} \right] - \text{erf} \left[ \frac{c - g}{\sqrt{2\sigma}} \right] > \text{erf} \left[ \frac{\sigma + g}{\sqrt{2\sigma^2}} \right] - \text{erf} \left[ \frac{c}{\sqrt{2}} \right]. \tag{A2}
\]
Thus, (A2) is shown for \( 0 < g \leq \sigma \). Next we consider the case \( g > \sigma \). Because the inequity \( e^{-(t-c/\sqrt{2})^2} > e^{-t^2} \) holds for \( t \geq c/\sqrt{2} \), we obtain
\[
\int_{-\infty}^{\infty} e^{-\left(\frac{t-c}{\sqrt{2\sigma}}\right)^2} dt > \int_{-\infty}^{\infty} e^{-t^2} dt
\]
\[
\Rightarrow \int_{-\infty}^{\infty} e^{-t^2} dt > \int_{-\infty}^{\infty} e^{-\left(\frac{t-c}{\sqrt{2\sigma}}\right)^2} dt
\]
\[
\Rightarrow \text{erf} \left[ \frac{c}{\sqrt{2}} \right] > \text{erf} \left[ \sqrt{2c} \right] - \text{erf} \left[ \frac{c}{\sqrt{2}} \right]. \tag{A3}
\]
Also we have \( e^{-(t-(\sqrt{2}c))^2} > e^{-t^2} \) for \( t \geq \sqrt{2}c \), and
\[
\int_{-\infty}^{\sqrt{2}c} e^{-\left(\frac{t-c}{\sqrt{2\sigma}}\right)^2} dt > \int_{-\infty}^{\sqrt{2}c} e^{-t^2} dt
\]
\[
\Rightarrow \int_{0}^{\sqrt{2}c} e^{-t^2} dt > \int_{-\infty}^{\sqrt{2}c} e^{-\left(\frac{t-c}{\sqrt{2\sigma}}\right)^2} dt
\]
\[
\Rightarrow - \text{erf} \left[ \frac{c - g}{\sqrt{2\sigma}} \right] > \text{erf} \left[ \frac{\sigma + g}{\sqrt{2\sigma^2}} \right] - \text{erf} \left[ \sqrt{2c} \right]. \tag{A4}
\]
Adding (A3) to (A4), we get
\[
\text{erf} \left[ \frac{c}{\sqrt{2}} \right] - \text{erf} \left[ \frac{c - g}{\sqrt{2\sigma}} \right] > \text{erf} \left[ \frac{\sigma + g}{\sqrt{2\sigma^2}} \right] - \text{erf} \left[ \sqrt{2c} \right]. \tag{A5}
\]
We have shown (A2) for \( 0 < g \leq \sigma \) and (A4) for \( g > \sigma \). Putting them together, we have (A1) for \( g > 0 \), and therefore Pr \( |E_2,ps \rangle \leq \text{Pr} \[ |E_2,np \rangle \) when \( \langle A \rangle \geq 1 \).

**Appendix B: Derivation of the \( z \) distribution function in Sec. IV**

A moment-generating function determines the distribution function of a random variable, and we can derive the distribution of \( z \) from its moment-generating function. Since the random variables \( x \) and \( y \) are independent, the moment-generating function of \( z = x + y \) satisfies
\[
E \left[ e^{\xi z} \right] = E \left[ e^{\xi x} \right] E \left[ e^{\xi y} \right], \tag{B1}
\]
where the E means the expectation value. It is known that the moment-generating function of the Gaussian distribution \( N(y) := e^{-y^2/(2\sigma^2)} / \sqrt{2\pi \sigma^2} \) is,
\[
E \left[ e^{\xi y} \right] = \int e^{\xi y} N(y) dy = e^{\frac{\xi^2}{2\sigma^2}}. \tag{B2}
\]
Then, the moment-generating function of the \( x \) distribution with postselection is
\[
E_{f_{ps}}[e^{\xi x}] = \int e^{\xi x} f_{ps}(x|g) dx
\]
\[
= \frac{1}{1 + |\langle \hat{A} \rangle_w|^2 + (1 - |\langle \hat{A} \rangle_w|^2)} e^{-\frac{\xi^2}{2\sigma^2}} \left[ (1 + |\langle \hat{A} \rangle_w|^2 + 2\text{Re}(\langle \hat{A} \rangle_w) e^{\sigma \xi} \right.
\]
\[
+ \left( 1 + |\langle \hat{A} \rangle_w|^2 - 2\text{Re}(\langle \hat{A} \rangle_w) e^{-\sigma \xi} \right] \tag{B3}
\]
and
\[
E_{f_{ps}}[e^{\xi z}] = E_{f_{ps}}[e^{\xi x}] E_{f_{ps}}[e^{\xi y}] = \frac{1}{1 + |\langle \hat{A} \rangle_w|^2 + (1 - |\langle \hat{A} \rangle_w|^2)} e^{-\frac{\xi^2}{2\sigma^2}} \left[ (1 + |\langle \hat{A} \rangle_w|^2 + 2\text{Re}(\langle \hat{A} \rangle_w) e^{\sigma \xi} \right.
\]
\[
+ \left( 1 + |\langle \hat{A} \rangle_w|^2 - 2\text{Re}(\langle \hat{A} \rangle_w) e^{-\sigma \xi} \right] \tag{B4}
\]
Because of the linearity of the expectation value, the distribution of \( z \) is given by
\[
f_{ps}(z|g) = \frac{1}{2\pi(\sigma^2 + s^2)} \left[ (1 + |\langle \hat{A} \rangle_w|^2 + 2\text{Re}(\langle \hat{A} \rangle_w) e^{-\frac{(\xi - 2s)^2}{2(\sigma^2 + s^2)}} \right.
\]
\[
+ \left( 1 + |\langle \hat{A} \rangle_w|^2 - 2\text{Re}(\langle \hat{A} \rangle_w) e^{\frac{(\xi - 2s)^2}{2(\sigma^2 + s^2)}} \right] \tag{B5}
\]
Similarly, we have the moment-generating function of the \( x \) distribution without postselection
\[
E_{f_{ps}}[e^{\xi z}] = e^{\frac{\xi^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi(\sigma^2 + s^2)}} e^{-\frac{(\xi - 2s)^2}{2(\sigma^2 + s^2)}} \tag{B6}
\]
and the distribution of \( z \)
\[
f_{np}(z|g) = \frac{|\langle i \rangle|^2}{\sqrt{2\pi(\sigma^2 + s^2)}} e^{-\frac{(\xi - 2s)^2}{2(\sigma^2 + s^2)}} + \frac{|\langle -i \rangle|^2}{\sqrt{2\pi(\sigma^2 + s^2)}} e^{-\frac{(\xi + 2s)^2}{2(\sigma^2 + s^2)}} \tag{B7}
\]
We have obtained (B3) and (B7), from which we can calculate the probabilities of the type 1 error Eq. (7) and those of the type 2 error Eqs. (8) and (9) in Sec. IV.

The distribution functions (B3) and (B7) can be derived by convolution, which is known to give a Gaussian channel. We can calculate the distribution function with postselection
by using for example

\[
\int \frac{e^{-\frac{(x-g)^2}{2\sigma^2}} - e^{-\frac{(x-g)^2}{2s^2}}}{\sqrt{2\pi}\sigma^2 + \sqrt{2\pi}s^2} \, dx
\]

\[
= \frac{1}{\sqrt{(2\pi\sigma^2)(2\pi s^2)}} \int e^{-\frac{(x^2 + (g - \mu)^2)}{2\sigma^2}} \left( e^{-\frac{(x^2 + (g - \mu)^2)}{2s^2}} \right)^2 \, dx
\]

\[
= \frac{1}{\sqrt{2\pi(\sigma^2 + s^2)}} \frac{2s^2 \sigma^2 e^{-\frac{(\mu - g)^2}{2(\sigma^2 + s^2)}}}{(2\pi\sigma^2)(2\pi s^2)(\sigma^2 + s^2)}
\]

\[
= \frac{1}{\sqrt{2\pi(\sigma^2 + s^2)}} e^{-\frac{(\mu - g)^2}{2(\sigma^2 + s^2)}}.
\] (B9)

We immediately see that Eq. (B8) equals to Eq (B5) and that a Gaussian channel and an additive white Gaussian noise are identical.

Appendix C: Evaluation of the test including the loss by postselection using the Lagrange multiplier method

In Sec. [11] we have considered the hypothesis testing if it was able to acquire the data at least once. Here, we consider the risk that the data cannot be obtained by failure of the postselection and discuss it by including it in the cost function (B3). To obtain the optimized function by the Lagrange multiplier method.

In order to treat the data loss by failure of the postselection taking the test (B13) into account, we propose a revised test by the following decision function;

\[
d(x) = \begin{cases} 
1 & \text{if } (f + |x|) > c_f \sigma \\
0 & \text{if } (f + |x|) < c_f \sigma \\
r & \text{otherwise}
\end{cases}
\] (C1)

We denote \( f \) and \( \bar{f} \) as success and failure of the postselection, respectively. In this test, we use the postselection result and the measurement result of the probe as statistics. The critical points \( c_f \) and \( c_f \) differ depending on the result of postselection, success or failure.

Here, we calculate the probabilities of the two types of error. The probability of the type 1 error is

\[
\Pr[|E_1|] = \Pr[d = 1|g = 0]
\]

\[
= \Pr[f, |x| > c_f \sigma |g = 0| + \Pr[\bar{f}, |x| > c_f |g = 0|
\]

\[
= 1 - \left( \frac{c_f}{\sqrt{2}} \right) \sqrt{\frac{2\pi}{\sigma^2}} \cdot |\langle f |i \rangle|^2 + \left( \frac{c_f}{\sqrt{2}} \right) \cdot |\langle \bar{f} |i \rangle|^2,
\] (C2)

and the probability of the type 2 error is

\[
\Pr[|E_2|] = \Pr[d = 1|g \neq 0]
\]

\[
= \Pr[f, |x| < c_f \sigma |g \neq 0| + \Pr[\bar{f}, |x| < c_f |g \neq 0|
\]

\[
= \frac{1}{4} \left( \left( |\langle f |i \rangle|^2 + |\langle \bar{f} |i \rangle|^2 \right) \left( \frac{c_f + g}{\sqrt{2\pi} \sigma^2} \right) + \left( |\langle f |i \rangle|^2 + |\langle \bar{f} |i \rangle|^2 \right) \left( \frac{c_f + g}{\sqrt{2\pi} \sigma^2} \right) \right)
\]

\[
+ \frac{1}{4} \left( \left( |\langle f |i \rangle|^2 - |\langle \bar{f} |i \rangle|^2 \right) \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) + \left( |\langle f |i \rangle|^2 - |\langle \bar{f} |i \rangle|^2 \right) \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) \right)
\]

\[
= \frac{1}{2} \left( \left( |\langle f |i \rangle|^2 - |\langle \bar{f} |i \rangle|^2 \right) \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) + \left( |\langle f |i \rangle|^2 - |\langle \bar{f} |i \rangle|^2 \right) \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) \right) + \left( \frac{c_f + g}{\sqrt{2\pi} \sigma^2} \right)
\] (C3)

If \( c_f = c_f \), the probability of the type 1 error becomes

\[
\Pr[|E_1|] = 1 - \left( \frac{c_f}{\sqrt{2}} \right),
\]

and the probability of the type 2 error becomes

\[
\Pr[|E_2|] = \frac{1}{2} \left( \left( \frac{c_f + g}{\sqrt{2\pi} \sigma^2} \right) + \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) \right).
\] (C5)

The both are independent of the postselection result. Then, we can substantially simplify the treatment of the case without postselection when \( c_f = c_f \) by the decision function (C1).

Let us consider how we express the errors of the weak measurement in the decision function (C1). When the postselection fails in the weak measurement experiment, we cannot obtain a measurement result. So in such a case, we cannot distinguish whether there is the interaction or not. But for convenience, we simply presume that there would not be the interaction. Since we are interested in the detection of the interaction, the result of no interaction is meaningless. We can conveniently handle this situation by setting \( c_f = \infty \) in our decision function. If \( c_f \) were \( \infty \), the alternative hypothesis would be always rejected by the test (C1) when the postselection fails. Hence, the decision function (C1) would cover the case with or without postselection including the data loss by failure of the postselection.

Meanwhile, if there is the situation such that we want to detect “vanishment” of an interaction which usually exists, the treatment of \( c_f \) as stated above is inconsequent. Here we note that if the \( c_f \) and \( c_f \) take the other value as presented above, we cannot give an obvious interpretation what the experimental situation means. Thus, it is often difficult to find out the physical meaning of the optimization of the \( c_f \) and \( c_f \). While such problems are remaining, we try out this Lagrange multiplier method.

From here, we calculate the critical points and the initial and final states of the measured system which optimize the test (C1) by the Lagrange multiplier method. To optimize the probability of the type 2 error while keeping the probability of the type 1 error at the significance level \( \alpha \) which is an arbitrary constant, we set the Lagrangian as

\[
\mathcal{L}(p_1, p_2, c_f, c_f, \lambda) = \Pr[|E_2|] + \lambda(\Pr[|E_1|] - \alpha)
\]

\[
= 1 - \left( \frac{p_1}{2} + \frac{p_2}{2} - \left( \frac{c_f + g}{\sqrt{2\pi} \sigma^2} \right) + \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) \right)
\]

\[
= 1 - \left( \frac{p_1}{2} + \frac{p_2}{2} + \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) + \left( \frac{c_f + g}{\sqrt{2\pi} \sigma^2} \right) \right)
\]

\[
+ \lambda \left( \frac{1}{2} \left( \left( \frac{c_f - g}{\sqrt{2\pi} \sigma^2} \right) + \left( \frac{c_f + g}{\sqrt{2\pi} \sigma^2} \right) \right) - \alpha \right),
\] (C6)
where \( \lambda \) is the Lagrange multiplier and the constraint condition comes from the standard strategy of the hypothesis testing as described in Sec. [112]. To simplify the notations, we denote \( p_1 := |\langle f | i \rangle|^2 \) and \( p_2 := |\langle f | \tilde{A} | i \rangle|^2 \). Note that \( \tilde{A}^2 = 1 \) and \( 0 < p_1, p_2 < 1 \). Varying the Lagrangian \( L \) with respect to \( \lambda \), the constrain condition reappears as

\[
0 = \frac{\partial L}{\partial \lambda} = \frac{1}{2} \left[ p_1 \left( 1 - \operatorname{erf} \left( \frac{c_f}{\sqrt{2}} \right) \right) \right] - \alpha. \tag{C7}
\]

Then, varying the Lagrangian \( L \) with respect to \( p_1 \) and \( p_2 \), we get

\[
0 = \frac{\partial L}{\partial p_1} = \frac{1}{4} \left[ \operatorname{erf} \left( \frac{c_f - \sigma}{\sqrt{2} \sigma} \right) + \operatorname{erf} \left( \frac{c_f + \sigma}{\sqrt{2} \sigma} \right) \right] - \lambda \left( 2 e^{-\frac{c_f^2}{2\sigma^2}} - \lambda \right). \tag{C8}
\]

Form these equations, we have

\[
0 = \frac{\partial L}{\partial p_1} - \frac{\partial L}{\partial p_2} = \frac{1}{2} \left( \operatorname{erf} \left( \frac{c_f}{\sqrt{2}} \right) - \operatorname{erf} \left( \frac{c_f}{\sqrt{2}} \right) \right) \left( 2 e^{-\frac{c_f^2}{2\sigma^2}} - \lambda \right). \tag{C10}
\]

So, we require either or both of \( \lambda = 2 e^{-g^2/2a^2} \) and \( c_f = c_f \). Then, varying \( L \) with respect to \( c_f \) and \( c_f \) gives

\[
0 = \frac{\partial L}{\partial c_f} = \frac{1}{4} \left( p_1 + p_2 \right) \left( e^{-\frac{(c_f - \sigma)^2}{2\sigma^2}} + e^{-\frac{(c_f + \sigma)^2}{2\sigma^2}} \right) + 2 \left( p_1 - p_2 - \lambda p_1 \right) e^{-\frac{c_f^2}{2\sigma^2}} \tag{C11}
\]

and

\[
0 = \frac{\partial L}{\partial c_f} = \frac{1}{4} \left( 2 - p_1 - p_2 \right) \left( e^{-\frac{(c_f - \sigma)^2}{2\sigma^2}} + e^{-\frac{(c_f + \sigma)^2}{2\sigma^2}} \right) + 2 \left( p_1 - p_2 \right) \left( 1 - \lambda p_1 \right) e^{-\frac{c_f^2}{2\sigma^2}}. \tag{C12}
\]

Here, we consider the case \( c_f = c_f = c \), where Eqs. (C8) and (C9) are fulfilled. The constrain constant (C7) becomes

\[
0 = \frac{1}{2} \left( 1 - \operatorname{erf} \left( \frac{c}{\sqrt{2}} \right) \right) - \alpha. \tag{C13}
\]

Because the \( \alpha \) is a constant, the \( c \) is fixed. Next, from Eqs. (C11) and (C12) we get

\[
0 = \frac{e^{-\frac{c_f^2}{2\sigma^2}}}{4} \sqrt{\frac{\pi}{2}} \left[ p_1 \left( e^{-\frac{(c_f - \sigma)^2}{2\sigma^2}} + e^{-\frac{(c_f + \sigma)^2}{2\sigma^2}} \right) - 2\lambda \right], \tag{C14}
\]

and

\[
0 = \frac{e^{-\frac{c_f^2}{2\sigma^2}}}{4} \sqrt{\frac{\pi}{2}} \left[ 2 \left( e^{-\frac{(c_f - \sigma)^2}{2\sigma^2}} + e^{-\frac{(c_f + \sigma)^2}{2\sigma^2}} \right) - \lambda \right], \tag{C15}
\]

respectively. The sum of Eqs. (C14) and (C15) gives

\[
0 = \frac{1}{2} \sqrt{\frac{2}{\pi}} \left( e^{-\frac{c_f^2}{2\sigma^2}} + e^{-\frac{c_f^2}{2\sigma^2}} \right) - \lambda \tag{C16}
\]

Thus, we obtain

\[
\lambda = e^{-\frac{c_f^2}{2\sigma^2}} \left( e^{\frac{c_f^2}{2\sigma^2}} + e^{\frac{c_f^2}{2\sigma^2}} \right). \tag{C17}
\]

Substituting \( c_f = c_f = c \) and \( \lambda = e^{-g^2/2a^2} \left( e^{g^2/a} + e^{-g^2/a} \right) \) into Eqs. (C11) and (C12), we have

\[
0 = \frac{e^{-\frac{c_f^2}{2\sigma^2}}}{4} \sqrt{\frac{\pi}{2}} \left( p_1 - p_2 \right) \left( e^{-\frac{c_f^2}{2\sigma^2}} - e^{-\frac{c_f^2}{2\sigma^2}} \right). \tag{C18}
\]

Then, we can find that \( c_f = c_f = c = 0 \) or \( p_1 = p_2 \) is needed. If \( c = 0 \), we obtain \( \alpha = 1/2 \) and \( \lambda = 2e^{-g^2/2a^2} \) from Eqs. (C13) and (C17). Because the significance level \( \alpha \) is not always 1/2, the \( c = 0 \) is not consistent. Then \( c_f = c_f \neq 0 \) and \( p_1 = p_2 \) is a solution.

Next from Eq. (C10), we discuss the case \( \lambda = 2e^{-g^2/2a^2} \).

Because we have already studied the case that \( c_f = c_f \) is satisfied simultaneously, hereafter we assume \( c_f \neq c_f \). Substituting \( \lambda = 2e^{-g^2/2a^2} \) into Eqs. (C11) and (C12), we have

\[
0 = \frac{1}{4} \sqrt{\frac{2}{\pi}} \left( p_1 + p_2 \right) \left( e^{\frac{c_f^2}{2\sigma^2}} - e^{\frac{c_f^2}{2\sigma^2}} \right)^2 e^{-\frac{c_f^2}{2\sigma^2}} \tag{C19}
\]

and

\[
0 = \frac{1}{4} \sqrt{\frac{2}{\pi}} \left( 2 - p_1 - p_2 \right) \left( c_f - c_f \right)^2 e^{-\frac{c_f^2}{2\sigma^2}} \tag{C20}
\]

From Eqs. (C19) and (C20), we can find that we need either condition as follows: the condition such that \( c_f = 0 \) and \( p_1 = p_2 = 1 \), or the condition such that \( c_f = 0 \) and \( p_1 = p_2 = 0 \). In both cases, the constrain condition (C7) becomes \( \alpha = 1/2 \). As stated in above, the \( \alpha \) is not necessarily 1/2.

So the condition \( \lambda = 2e^{-g^2/2a^2} \) is not proper.

Therefore, we conclude the solution is \( c_f = c_f \neq 0 \) and \( p_1 = p_2 \). From \( p_1 = p_2 \), we derive

\[
0 = |\langle f | i \rangle|^2 - |\langle f | A | i \rangle|^2 = |i \rangle \left( |f \rangle - |A \rangle |f \rangle \right) i\rangle \tag{C21}
\]

\[
|f \rangle = A \langle f \rangle \tag{C21}
\]
or
\[ 0 = |\langle f| i \rangle|^2 - |\langle f| A|i \rangle|^2 = \langle f | (i - A|i \rangle A|i |f \rangle, \]
\[ \implies \pm |i \rangle = \bar{A}|i \rangle. \] (C22)

Then, \( p_1 = p_2 \) means that the preselected state \(|i \rangle\) or the postselected state \(|f \rangle\) equals to an eigenstate of \( \bar{A} \). As we have noted, the case \( c_j = c_f \) corresponds to the measurement without postselection. Thus, the result of the postselection has nothing to do with the test. The state of the postselection \(|f \rangle\) is not essential. Consequently, the optimal condition for the test \((\text{C1})\) is that the preselected state is an eigenstate of the measured observable \( A \) and that we do not postselect. We caution the readers that this Appendix gives a nothing but a crude trial. We also note that the Lagrange multiplier method gives the stationary point to the utmost, and they might not be the minimum.

[1] K. J. Resch, Science 319, 733 (2008).
[2] A. M. Steinberg, Nature 463, 890 (2010).
[3] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. 60, 1351 (1988).
[4] Y. Aharonov and A. Botero, Phys. Rev. A. 72, 052111 (2005).
[5] Y. Aharonov and D. Rohrlich, Quantum Paradoxes (Wiley-VCH, Weinheim, 2005).
[6] Y. Aharonov and L. Vaidman, in Time in Quantum Mechanics, edited by J.G. Muga et al. (Springer, Berlin Heidelberg, 2008), 399.
[7] R. Jozsa, Phys. Rev. A 76, 044103 (2007).
[8] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, Phys. Rev. D 40, 2112 (1989).
[9] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Rev. Mod. Phys. 86, 307 (2014).
[10] S. Wu and Y. Li, Phys. Rev. A 83, 052106 (2011).
[11] X. Zhu, Y. Zhang, S. Pang, C. Qiao, Q. Liu, and S. Wu, Phys. Rev. A 84, 052111 (2012).
[12] T. Koike and S. Tanaka, Phys. Rev. A 84, 062106 (2011).
[13] K. Nakamura, A. Nishizawa, and M.-K. Fujimoto, Phys. Rev. A 85, 012113 (2012).
[14] Y. Susa, Y. Shikano, and A. Hosoya, Phys. Rev. A 85, 052110 (2012).
[15] A. Di Lorenzo, Phys. Rev. A 87, 046101 (2013).
[16] Y. Susa, Y. Shikano, and A. Hosoya, Phys. Rev. A 87, 046102 (2013).
[17] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Phys. Rev. Lett. 66, 1107 (1991).
[18] O. Hosten and P. Kwiat, Science 319, 787 (2008).
[19] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Phys. Rev. Lett. 102, 173601 (2009).
[20] G. I. Viza, J. Martines-Rincon, G. A. Howland, H. Frostig, I. Shomroni, B. Dayan, and J. C. Howell, Opt. Lett. 38, 2949-2952 (2013).
[21] A. Nishizawa, K. Nakamura, and M.-K. Fujimoto, Phys. Rev. A 85, 062108 (2012).
[22] A. N. Jordan, J. Martines-Rincon, and J. C. Howell, Phys. Rev. X 4, 011031 (2014).
[23] J. Lee and I. Tsutsumi, Quantum Studies: Mathematics and Foundations 1 pp 65-78, (2014).
[24] G. C. Knee and E. M. Gauger, Phys. Rev. X 4, 011032 (2014).
[25] G. C. Knee, G. A. D. Briggs, S. C. Benjamin, and E. M. Gauger, Phys. Rev. A 87, 012115 (2013).
[26] S. Tanaka and N. Yamamoto, Phys. Rev. A, 88, 042116 (2013).
[27] C. Ferrie and J. Combes, Phys. Rev. Lett. 112, 040406 (2014).
[28] B. Roy Frieden, Science from Fisher Information: A Unification (Cambridge university press, Cambridge, 2004).
[29] G. C. Knee, J. Combes, C. Ferrie, and E. M. Gauger, arXiv:1410.6222 (2014).
[30] L. Vaidman, arXiv:1402.0199 (2014).
[31] Y. Kedem, arXiv:1402.1352 (2014).
[32] C. Ferrie and J. Combes, arXiv:1402.2954 (2014).
[33] J. Neyman and E. S. Pearson, Egon S, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 231 (1933), 694-706.
[34] T. S. Ferguson, Mathematical Statistics : a decision theoretic approach, (New York, Academic Press, 1967).
[35] C. R. Rao, Linear Statistical Inference and its Applications (2nd ed, New York, Chichester, Wiley, 1973).
[36] J. C. Kiefer, Introduction to statistical inference, (New York, Springer, 1987).
[37] J. Pfanzagl and R. Hamböker, Parametric statistical theory, (Berlin, W. de Gruyter, 1994).
[38] H. Nagaoka, Proc. Int. Symp. on Inform. Theory 198, 12489-12504 (2006).
[39] A. Fujiwara, J. Phys. A: Math. Gen, 39, 12489-12504 (2006).
[40] In a randomized test, we randomly accept the null hypothesis with the probability \( r \). This test is widely used at the critical point which is not included in neither the accept nor the rejection region. For the distribution function of continuous variable, the probability for the continuous random variable coincides with the critical point is zero. This slightly modified test does not change our conclusion.
[41] E. L. Lehmann, Testing statistical hypotheses (New York : Wiley, 1959).
[42] A. C. Elitzur and L. Vaidman, Found. Phys. 23, 987 (1993).
[43] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M. A. Kasevich, Phys. Rev. Lett. 74, 4763 (1995).
[44] J. O. Berger, Statistical decision theory and Bayesian analysis (New York, Springer, 1985), 20-21; J. O. Berger and M. Delampady, Statist. Sci. Vol. 2, Number 3 (1987), 317-335.
[45] A cost function represents the cost or risk that we have to pay in the process which we want to evaluate. We can choose the cost function as we like.