論文 / 著書情報
Article / Book Information

題目(和文)	弱値増幅における最適プローブと統計的検定
Title(English)	Optimal Probe and Statistical Testing in Weak-Value Amplification
著者(和文)	須佐友紀
Author(English)	Yuki Susa
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第10052号, 授与年月日:2016年3月26日, 学位の種別:課程博士, 審査員:宗宮 健太郎,旭 耕一郎,西森 秀稔,斎藤 晋,上妻 幹旺,細谷 暁夫,鹿野 豊
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第10052号, Conferred date:2016/3/26, Degree Type:Course doctor, Examiner:,,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Doctoral Thesis

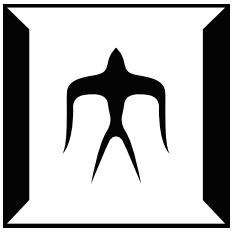
Optimal Probe and Statistical Testing in Weak-Value Amplification

YUKI SUSA

A thesis submitted in fulfillment for
the degree of Doctor of Science

in the

Department of Physics
Tokyo Institute of Technology



February 2016

Abstract

Recently, the weak-value amplification (WVA) has been extensively investigated as a promising quantum technique for improving an accuracy of a precision measurement. The basic idea of WVA comes from the weak measurement introduced by Aharonov, Albert, and Vaidman in 1988. The weak measurement is the indirect quantum measurement with the weak interaction and postselection of the final state of the measured system. The attention point is that we can measure the quantity called the weak value, which appears in the shift of the measuring probe given by the weak measurement. Because the weak value can be outside the range of the eigenvalues, the shift of the measuring probe obtained by the weak measurement becomes larger than the one given by the conventional measurement. This shift amplification effect by the weak value is the WVA, which has been studied for precisely measuring the small coupling constant.

In this thesis, we propose the two possible applications of the WVA. First, we analytically obtain the optimal probe wave function for a given weak value by the variational method. It is shown that the amplification factor obtained by the optimal probe has no-upper bound while the one in the Gaussian probe already reported by some researchers has the upper bound. Furthermore, the variance of the optimal probe after the measurement can be zero. We can derive the optimal probe by the Lagrange multiplier method for maximizing the shift of the probe position or minimizing the variance of the final distribution.

Second we have developed the way which has a technical advantage of the WVA from the viewpoint of statistics. It is often argued that the WVA is not helpfull for the parameter estimation. Therefore we consider the statistical inferences that the WVA will be advantageous. We propose a method to determine whether the interaction is present or not and evaluate the capability of the WVA for this interaction detection problem by using the hypothesis testing in statistics. As the main result, it is shown that the merit of the WVA is the reduction of the possibility to miss the presence of the interaction more than the conventional measurement while keeping the probability of a misdetection, when the absolute value of the weak value is outside the range of the eigenvalues. In this discussion, we assume that the number of the obtained data is infinitely large and we neglect the data loss by the failure of the postseleciton.

Acknowledgments

I am deeply grateful to my supervisor, Prof. Kentaro Somiya for allowing me to study the research theme that I hope and constant encouragement. I would like to express the deepest appreciation to my ex-supervisor, Prof. Akio Hosoya for incisive comments, constructive advices and long-term support since my undergraduate study.

I would like to offer my special thank to my co-workers, Prof. Yutaka Shikano at Institute of Molecule Science, and Saki Tanaka at Keio University for valuable discussion to solve interesting problems, the optimal probe wave function and the statistical hypothesis testing in the weak-value amplification, respectively. I thank Prof. Antonio Di Lorenzo at Universidade Federal de Uberlandia for valuable discussion about the optimal probe wave function, and Prof. Fuyuhiko Tanaka at Osaka University for comment on the statistical inference. I appreciate helpful comments on my research from Prof. Masa-Katsu Fujimoto and Kouji Nakamura at National Astronomical Observatory of Japan, Atsushi Nishizawa at California Institute of Technology (Caltech), and Iinuma Masataka at Hiroshima University. I would like to thank to Prof. Yanbei Chen and Haixing Miao for my visiting Caltech and discussion through MQM telecon.

I also thank previous and current Somiya lab members, Shinichiro Ueda, Nana Saito, Yukihide Sakihama, Jumpei Kato, Ayaka Kumeta, Meriem Boubdir, Kazushiro Yano, Sho Atsuta, Yuu Kataoka, Junko Kasuya, Takuya Yaginuma, and Tjark Miener for pleasant lab life.

I thank the financial supports from JSPS Research Fellowships for Young Scientists (Grant No. 25008633) and the Global Center of Excellence Program "Nanoscience and Quantum Physics" at Tokyo Institute of Technology that made it possible to complete this study during my doctoral course.

Finally, I would like to express my gratitude to my family for their support and encouragement.

List of Publications

1. Yuki Susa, Yutaka Shikano, and Akio Hosoya
   "Optimal probe wave function of weak-value amplification"
   Physical Review A **85**, 052110 (2012), doi:10.1103/PhysRevA.85.052110
   e-Print: arXiv:1203.0827

2. Yuki Susa, Yutaka Shikano, and Akio Hosoya
   "Reply to "Comment on 'Optimal probe wave function of weak-value amplification' " "
   Physical Review A **87**, 046102 (2013), doi:10.1103/PhysRevA.87.046102
   e-Print: arXiv:1304.1352

3. Yuki Susa, and Saki Tanaka
   "Statistical Hypothesis Testing by Weak-Value Amplification: Proposal and Evaluation"
   Physical Review A **92**, 012112 (2015), doi:10.1103/PhysRevA.92.012112
   e-Print: arXiv:1502.06334

4. Yuki Susa
   "Physical description of statistical hypothesis testing for a weak-value amplification experiment using a birefringent crystal"
   Physical Review A **92**, 022118 (2015), doi:10.1103/PhysRevA.92.022118
   e-Print: arXiv:1505.07199

Contents

Chapter 1

Introduction

1.1 Background

It is undoubted that a progress of a measurement technique substantially contributes to developing physics. As a recent topic, for example, the alternative uncertain principles [1–3] instead of Heisenberg's one [4] were theoretically proposed and some researchers experimentally demonstrated by the neutron spin [5] and the optical experiments [6]. In other instances, gravitational wave detectors, which are large Michelson interferometers, have been developed for an ultra-high-precision measurement [7, 8]. If the detectors observe a gravitational wave, we can make a definite progress in the researches of general relativity, cosmology, and astronomy. Furthermore, because the quantum uncertainty principle limits the sensitivity of the detector [9, 10], we need to study quantum measurement, quantum optics, and optomechanics [11–13], which will also contribute to a further measurement technique.

Meanwhile, in recent years, the *weak-value amplification* (WVA) has been intensively developed as a promising quantum measurment technique for improving an accuracy of a precision measurement [14–16]. The basic idea of the WVA emerged from the *weak measurement* introduced by Aharonov, Albert, and Vaidman in 1988 [17]. The weak measurement is described as an indirect quantum measurement with postselection of the final state of the measured system and has been studied by many quantum physicists in recent years [18]. For example, the experiment in Ref. [6] used the weak measurement. Usually, we assume that the interaction Hamiltonian is of the von Neumann type which gives rise to a translation of the probe position distribution. The interesting point is that

we can measure the quantity called the *weak value* by the weak measurement. The weak value is defined as

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

where $|i\rangle$ and $|f\rangle$ are the pre- and postselected states and $\hat{A}$ is an observable in the measured system, respectively. The weak value appears in the shift of an expectation value of the measuring probe induced by the interaction between the measured system and the probe. When the coupling is very weak, the shift is proportional to the weak value and the coupling constant [17]. The term of *weak* in the weak measurement comes from the weak coupling in the original proposal. On the other hand, the shift given by the conventional indirect measurement is the eigenvalues of the measured system observable and the coupling constant. We focus on the property of the weak value, that the weak value can be outside the range of the eigenvalues. Hence the shift of the measuring probe obtained by the weak measurement can be larger than the one given by the conventional measurement [19]. We call this shift amplification effect by the weak measurement as the weak-value amplification and some researchers have extensively studied the WVA for the purpose of measuring the small coupling constant [18].

Actually, the amplification effect has been demonstrated in various experiments [20–30]. Especially, an experiment for measuring a beam displacement shows a significant amplification effect. Setting a single birefringent cystal between the two poralizers, Ritchie *et al.* monitored the large beam split by a crystal with tuned poralizers [20]. We see this experiment in greater detail in Chap. 6. Hosten and Kwiat succeeded in the first observation of the spin Hall effect of light by the WVA [21]. The WVA in the Sagnac interferometer was done by Dixon *et al.* [22]. They measured the beam deflection caused by the tilted mirror with postselection, and they accomplished about 80 times amplification of the shift of the beam axis by the mirror. The WVA was also used for the velocity measurement by Viza *et al.* with the Michelson interferometer and measured the velocity of the longitudinal moving mirror from the photons at the output port. They have achieved the measurement of 400 fm/s [28].

We have some theoretical discussions on the properties of the WVA. Although the original proposal by AAV [17] used the weak coupling approximation, Wu and Li considered the higher-order terms in the coupling constant and showed the bound of the amplification factor [31]. The amplification limit is analytically given by full-order caluculation on the assumption that the measured system is a two-state system [32–34]. We note

that people often assume that the initial probe wave function has a Gaussian profile. Actually, Ref. [35] claimed that the amplification limits essentially depends on the initial probe state and the number of the distinct eigenvalues of the measured system.

Some researchers have theoretically studied a technical utility of the WVA. In Ref. [36], Nishizawa *et al.* have compared the amplified shift and the shot noise in an optical interferometer. Also the same resercher has studied the WVA in the Michelson interferometer which has two kinds of the fundamental quantum noise, the shot noise and the radiation-pressure noise [37]. Jordan *et al.* have analyzed the several types of the technical noise to conclude that the WVA has practical advantage for a precision measurement [38]. Lee and Tsutsui have evaluated the three principal noises, which show up in a standard measurement model [39]. They have shown that the WVA has advantage for the observation of interaction.

There is the trade-off that the larger the amplification factor is, the smaller the number of available measurement data becomes due to the postselected state almost orthogonal to the preselected state. Some researchers worry that this trade-off might be considered as a disadvantage of the WVA for a precision measurement usually requiring a large number of data [40–44]. They had discussed this issue by the estimation theory, which is one of the statistical inferences. The purpose of this discussion is the estimation of the value of the coupling constant which indicates the interaction strength between the measured system and the measuring probe. They have evaluated the estimation accuracy of the coupling constant with the weak measurement and also with the conventional measurement. It is well known that the inverse of the Fisher information multiplied by the number of obtained data gives the lower bound of the mean squared error of the estimator, i.e., the Cramér-Rao inequality [45]. Applying this bound to both the weak measurement and the conventional one, they concluded that the weak one is inferior to the conventional one for the parameter estimation due to the data loss by the failure of the postselection. Some researchers mention that the data loss by the postselection dose not need to be considered in practical cases [32]. Also the researchers theoretically proposed ways to circumvent this weak point of the postselection by recycling photons Refs. [46, 47] or by entanglement [48, 49]. Actually, there remain controversies over this discussion whether the WVA is usefull or not [30, 38, 50–52].

1.2 Aim of this thesis

As we will see in the subsequent sections the original idea of the WVA by AAV may not have an advantage over the conventional measurement. One might contemplate various ways of its modifications to make the WVA advantageous; choose the probe wave function other than the standard Gaussian function or even change the task of measurement itself and probably more.

In this thesis, we concentrate on the two applications of the WVA from the different viewpoints, a probe engineering to give a large amplification and a task of the statistical hypothesis testing for the interaction detection problem rather than the coupling parameter estimation. At the moment other possibilities cannot be excluded but we believe the present two are most promising, especially the second one.

First, as we have introduced in the previous section, Ref. [35] has shown that the amplification limit depends on the initial probe state. They have also calculated the amplification limit by using a variational method. Especially, they have exemplified the Stern-Gerlach experiment, the measured system of which is the two-state system, and suggested that we can enlarge the maximal shift of the position expectation value by broadening the initial probe wave function. We note that they have used the assumption that the interaction strength is very weak, i.e., $g \ll 1$ for the von Neumann Hamiltonian. Meanwhile, according to Refs. [32–34], we can calculate the expectation value of the probe position and momentum without any approximation when the measured system is the two-state system. Accordingly, we will evaluate the maximal shift of the probe position by the full-order calculation for the two-state system. The Gaussian probe case has the upper bound of the maximal shift. Here, the question is raised whether any initial probe wave function has the upper bound or not. In this thesis, we analytically derive the optimal probe wave function, which maximizes the shift of the probe, by the Lagrange multiplier method. It is shown that the amplification factor given by the optimal probe has no upper bound, while the amplification factor in the Gaussian probe already reported by some researchers has the upper bound for any weak values [32–34]. We also show that the optimal probe wave function is derived by the variational method for minimizing the variance of the final distribution. The optimal probe wave function is the unique solution of the two variational problems. For a specific weak value, the final distribution can be described as the Kronecker delta, the variance of which is zero. To produce this probe, we need to know the weak value, which can be calculated from an experimental setup, and the value of the coupling constant, which is an unknown parameter that we want

to measure. It is a practical strategy that we use the reasonably guessed value of the coupling constant and repeatedly update the value from the measurement results. Our result suggests that varying the initial probe will give benefit for developing the WVA or a precise measurement.

Next, we have developed a way which has a technical advantage of the WVA from the viewpoint of statistics. As we have noted above, some researchers worry that the failure of the postselection makes the number of available measurement data small, which might be a disadvantage of the WVA for a precisely estimating parameter usually requiring a large number of data. Actually, we see it with the typical example in Sec. 2.4 that the weak measurement is worse for the parameter estimation than the conventional measurement due to the failure of the postselection. Additionally, as stated in Sec. 2.4, we find that the Fisher information with postselection orthogonal to preselection is smaller than one with postselection parallel to preselection regardless of whether the data loss is considered or not. When postselecion is orthogonal to preselection, the weak measurement gives large a weak value and large amplification. Basically, the amplification by the weak measurement is not helpful for estimating regardless of the failure of the postselection. Thereupon we consider a statistical inference problem with the exception of the estimation method, in which the amplification effect would be beneficial even only the case that the number of the obtained date is infinitely large.

Here we study the detection problem that we determine whether the interaction is present or not between the measured system and the measuring probe. In this problem, we evaluate the distinguishing capability of the weak measurement and the conventional measurement and compare them. To see such a problem, the hypothesis testing method is usually used [53–55]. The similar detection problem has been treated also in Refs. [39, 42]. The authors of Ref. [39] did not used the hypothesis testing method. They supposed that the interaction is present if the expectation value of the final probe is non-zero for a zero-mean initial probe. They compared the expectation value with the noises (the systematical noise , the statistical noise, and the approximation error) to obviously judge whether the expectation value is non-zero or not. Further they find a region of the weak value that the weak measurement has the merit to find the non-zero expectation value. Here we remark that the region depends on the coupling constant that we want to measure. Therefore we need to know the coupling constant to configure an experiential setup. Additionally , in this method, we will miss the case that the interaction is present while the expectation value is zero. The authors of Ref. [42] evaluated the detection capability of the weak measurement with the likelihood-ratio test method, which is widely

used in the hypothesis testing for a practical reason. As mentioned later, the likelihood-ratio test is not always appropriate. We have another proper method to consider the detection problem. Therefore, we reconsider this problem properly using the standard hypothesis testing method to obtain a definite result.

In the hypothesis test, we set up the two contradictory hypotheses, a null hypothesis and an alternative hypothesis, and determine which one is more likely on the basis of the experimental results. We apply the hypothesis testing for the weak measurement and propose the appropriate test function for a particular testing problem that we determine whether or not the interaction is present. Namely, for the coupling constant g, we set the null hypothesis that the interaction is absent, i.e., $g = 0$, and the alternative hypothesis represents that the interaction is present, i.e., $g \neq 0$. Such a testing problem is classified as the two-sided test in the hypothesis testing theory. We have to choose an appropriate test function, which is the criterion of determination, prior to testing. In the two-sided test, a test function giving the uniformly most powerful unbiased (UMPU) test is most appropriate, the detection power of which is grater than one given by any other test functions and the probability of the misdection is a significance level. We remark that, in one-sided test, the uniformly most powerful (UMP) test is most appropriate, a detection power of which is grater than one given by any other test functions and the probability of the misdection is lower than a significance level. Often, the likelihood-ratio test is used for the testing problem such that the both hypotheses are simple, in which the test is UMP. In the likelihood-ratio test, comparing the two likelihood functions, i.e., probability distributions obtained by the measurement, one given in the null hypothesis and the other given in the alternative hypothesis, we determine which hypothesis is more likely. Noting that, in problem of the two-sided test, the UMP test does not exist. Although the likelihood-ratio test is usually used for practical reasons, e.g. Ref. [42], the likelihood-ratio test is not appropriate for our testing problem described as the two-sided test. Thus we should consider a test function giving the UMPU test for the two-sided test.

As the main result, it is shown that the merit of the WVA is the reduction of the possibility to miss the presence of the interaction more than the conventional measurement while keeping the probability of a misdetection, when the absolute value of the weak value is outside the range of the eigenvalues. In this discussion, we assume that the number of the obtained data is infinitely large and we neglect the data loss by the failure of the postseleciton. Additionally, we show that our proposed test function can be a UMPU test or a UMP test, which gives a statistically good test. It is also shown that our result

remains valid under an additive white Gaussian noise typically occurred in an electric circuit or device.

Furthermore, we apply our proposed testing method to the famous weak measurement experiment demonstrated by Ritchie *et al.* [20] to get the physical intuition of the proposed testing method. This experiment uses the two polarizers and the single birefringent crystal, which is originally designed for measurement of a weak value. In this framework, we consider the testing problem to determine whether a crystal used is birefringent or not. We show that the case of the almost orthogonal angles of the two polarizers, which does not satisfy the weak coupling approximation, make the WVA notable powerful in terms of the testing power. This result will enhance the physical understanding of the WVA and of the proposed testing method.

1.3 Organization

In Chap. 2, we review the concept of the weak measurement and the WVA following the original proposal by Aharonov, Albert, and Vaidman [17]. Also we see the weak-value amplification in a full-order calculation with the assumption that the measured system is two-quantum state system described as the Bloch sphere. We see the issue from the viewpoint of the estimation theory that the failure of the postselection causes the disadvantage of the WVA. By calculating the classical Fisher information, we can evaluate the estimation capability of the weak measurement.

In Chap. 3, we explicitly derive the optimal probe wave function in the momentum space by the Lagrange multiplier method in two ways: maximizing the shift of the probe position and minimizing the position variance of the final probe distribution. This chapter is based on Refs. [58, 59].

In Chap. 4, we recapitulate the standard concept of the statistical hypothesis testing for the subsequent chapter. The purpose of the hypothesis testing method is to determine the right hypothesis. We explain two types of errors that disturb the determination and quantitatively define the testing power. We introduce the test function, which gives a criterion of the determination. In the hypothesis testing, we have prescriptions for preparing the appropriate test function in response to a testing problem.

In Chap. 5, we consider the testing problem that we determine whether the interaction is present or not, which is treated as the two-sided test. We propose the test function in

which the interaction is supposed to be present when the variance of the measurement data deviates from the initial probe fluctuation. With this test function, we evaluate the testing powers of the weak measurement and the conventional measurement. This chapter is based on Ref. [60].

In Chap. 6, we see the physical intuition of the hypothesis testing method proposed in the previous chapter through the famous optical experiment for measuring the weak value demonstrated by Ritchie *et al.* [20]. We review the experiment which uses the two polarizers and a single birefringent crystal. In this thesis, we regard that this experiment is used for distinguishing whether the crystal used is birefringent or not. We find that case that postselected state completely orthogonal to the preselected one gives the striking amplification effect and the notable testing power, where the weak coupling approximation breaks down. This chapter is based on Ref. [61].

We give the concluding remarks of this thesis in Chap. 7. Especially, we summarize the discussion with the focus on the hypothesis testing with the weak measurement.

Subsequently, we note some technical matters in the Appendices. In Appendix A, we show an example of the position expectation value of the final probe in the full-order calculation. Here we do not restrict our attention to the case that the initial probe is Gaussian.

In Appendix B, we derive the Cramér-Rao Inequality and calculate the classical Fisher information for the weak measurement and the conventional measurement. We use them in Sec. 2.4.

In Appendix C, we review the optical shot noise in the weak measurement, which is one of the fundamental noises in an optical experiment, referring to Ref. [36]. We see that the shot noise is always larger than the variance of the final probe.

In Appendix D, we consider two other situations of the testing problem. In Appendix D.2, we consider the testing problem with a small interval null hypothesis such as $|g| \leq \varepsilon$ for small ε. This discussion is motivated by Refs. [56, 57], the author of which mentions that one might care a point null hypothesis such as $g = 0$. Also, in Appendix D.3, we secondarily discuss the testing including the risk of the data loss by failure of the postselection. For this discussion, we alternatively propose the makeshift test function which has a few defects in its physical interpretation. This section is based on Appendix C in Ref. [60].

Chapter 2

Basic Concept of Weak-Value Amplification

2.1 Preface

In this chapter, we introduce the basic concept of the weak-value amplification (WVA). The WVA is derived from the weak measurement proposed by Aharonov, Albert and Vaidman (AAV) [17]. At the stage of their proposal, they assumed that the coupling constant is small. We can find that the strange quantity called the weak value in the shift of probe position or momentum, which brings an amplification effect. Recently, some researchers have analyzed the weak measurement without any approximation under the particular assumption [32–34]. They have shown that the amplification factor has an upper bound and we can see that the factor depends on the initial measuring probe state. Also there is a discussion from the viewpoint of the estimation theory which is one of the statistical inferences [41].

2.2 Weak Measurement in AAV Formalism

We recapitulate the weak measurement and the conventional measurement. These measurements are described as indirect quantum measurements, which need the measured system and the measuring probe. Initially, we prepare the preselected state $|i\rangle$ of the measured system and the initial state $|\psi\rangle$ of the measuring probe.

First, we take the Gaussian profile which has the width σ for the initial state of the measuring probe, the wave function of which in the position x space is

$$\psi(x) := \langle x|\psi\rangle = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{4}} e^{-\frac{x^2}{4\sigma^2}}. \tag{2.1}$$

With the Fourier transform, we can derive the initial probe wave function in the momentum space p as

$$\tilde{\psi}(p) := \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi}} \int dx\ e^{-ipx}\langle x|\psi\rangle = \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{4}} e^{-\sigma^2 p^2}. \tag{2.2}$$

We note that the expectation values of the position $\langle\hat{x}\rangle$ and the momentum $\langle\hat{p}\rangle$ given by the probe state $|\psi\rangle$ are both 0. We have an interaction between the measured system and the measuring probe. Usually, we assume that the interaction Hamiltonian is von Neumann type described as

$$\hat{H}_{\text{int}} = g\delta(t)\hat{A} \otimes \hat{p}, \tag{2.3}$$

where g is a coupling constant, $\hat{A}$ is an observable in the measured system and $\hat{p}$ is the momentum operator of the measuring probe which is conjugate to the position operator $\hat{x}$. For simplicity, we assume that the probe and system momentarily interact at $t = 0$. Therefore the unitary operator is given by $\hat{U} = \exp(-ig\hat{A} \otimes \hat{p})$ which represents time-evolution. This interaction gives the position displacement to the measuring probe wave function. For example, we assume that the measured system is a two-state system consisting of the states $|+\rangle$ and $|-\rangle$ and the observable operator $\hat{A}$ is given as

$$\hat{A} = \lambda_+|+\rangle\langle+| + \lambda_-|-\rangle\langle-|, \tag{2.4}$$

where $\lambda_\pm \in \mathbb{R}$ are the eigenvalues of the states $|\pm\rangle$, respectively. After the interaction, the probe probability distribution in the position space becomes

$$\begin{aligned} f_c(x|g) &= |\langle x|\psi_c\rangle|^2 = |\langle x|\hat{U}|\psi\rangle|i\rangle|^2 \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \left(|\langle+|i\rangle|^2 e^{-\frac{(x-\lambda_+ g)^2}{2\sigma^2}} + |\langle-|i\rangle|^2 e^{-\frac{(x-\lambda_- g)^2}{2\sigma^2}} \right). \end{aligned} \tag{2.5}$$

We can see that the probability distribution of the probe position is shifted to two split Gaussian distributions, the peak positions of which are (the interaction strength g) $\times$ (the eigenvalues $\lambda_\pm$). Here we focus on the coefficients $|\langle+|i\rangle|^2$ and $|\langle-|i\rangle|^2$ of the two Gaussian distributions. If the initial state is $|i\rangle = |+\rangle$, for example, the final probe distribution

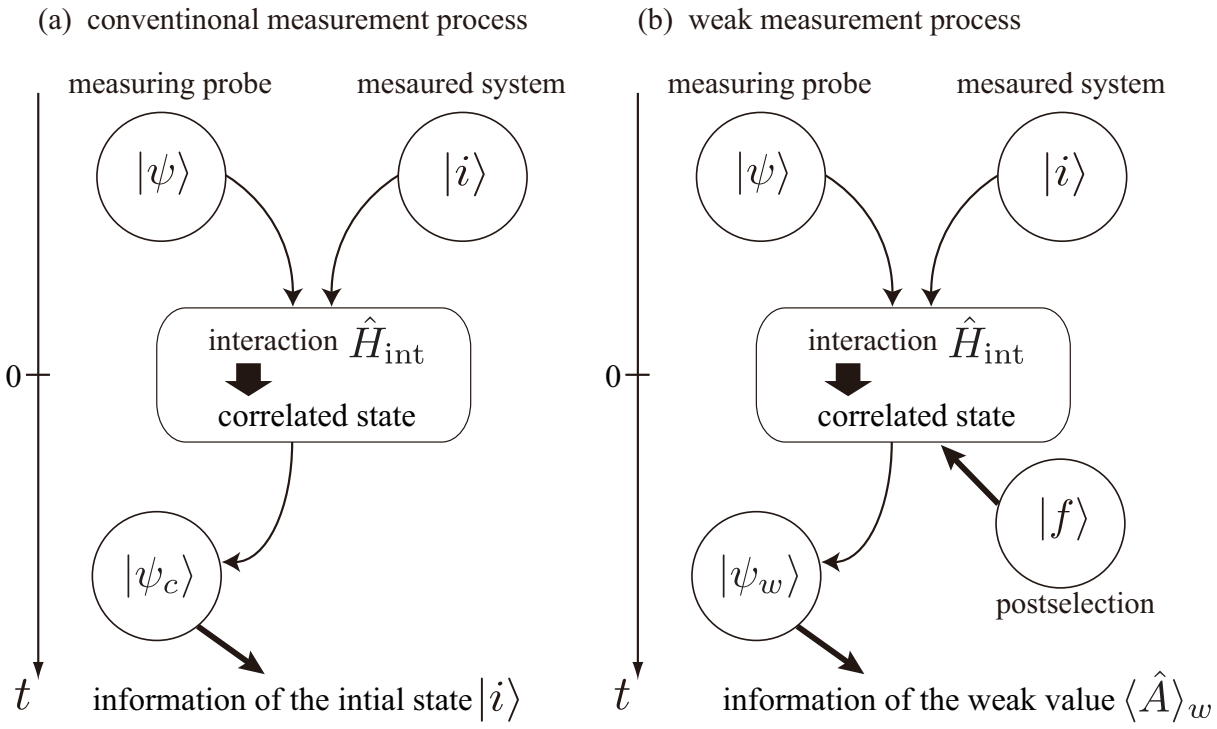


Figure 2.1: Schematic diagrams of (a) the conventional measurement process and (b) the weak-measurement process.

becomes a single Gaussian, the peak position of which is $g\lambda_+$ and so on. Thus we can get information of the initial state $|i\rangle$ from the final probe distribution $f_c(x|g)$. This is the standard process of the conventional measurement [Fig. 2.1(a)].

On the other hand, the weak measurement is originally proposed as the method to extract the weak value of the measured system [Fig. 2.1(b)]. In the weak measurement processes, we select the final state $|f\rangle$ of the measured system. Performing the postselection is a key distinct from the conventional measurement. By the postselection, the final probe distribution is significantly changed from the one given by the conventional measurement in general. AAV [17] evaluated the final probe state in the weak coupling approximation $g \ll 1$ as

$$|\psi_w\rangle = \langle f|e^{-ig\hat{A}\otimes\hat{p}}|i\rangle|\psi\rangle \approx \langle f|(1-ig\hat{A}\otimes\hat{p})|i\rangle|\psi\rangle = \langle f|i\rangle\left[1-ig\langle\hat{A}\rangle_w\hat{p}\right]|\psi\rangle, \tag{2.6}$$

where the weak value is defined as

$$\langle\hat{A}\rangle_w := \frac{\langle f|\hat{A}|i\rangle}{\langle f|i\rangle} \in \mathbb{C}. \tag{2.7}$$

The final probe wave function in the position space is given by

$$\begin{aligned}\langle x|\psi_w\rangle \approx & \langle f|i\rangle \left[1 - ig\langle\hat{A}\rangle_w \left(-i\frac{\partial}{\partial x}\right)\right] \langle x|\psi\rangle \\ = & \langle f|i\rangle \left[1 - g\langle\hat{A}\rangle_w \left(-\frac{x}{2\sigma^2}\right)\right] \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{4}} e^{-\frac{x^2}{4\sigma^2}} \\ \approx & \langle f|i\rangle e^{g\langle\hat{A}\rangle_w \frac{x}{2\sigma^2}} \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{4}} e^{-\frac{x^2}{4\sigma^2}} \\ = & \langle f|i\rangle \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{4}} e^{-\frac{1}{4\sigma^2}\left(x - g\mathrm{Re}\langle\hat{A}\rangle_w\right)^2} e^{\frac{g^2\mathrm{Re}\langle\hat{A}\rangle_w^2}{4\sigma^2} + ig\mathrm{Im}\langle\hat{A}\rangle_w \frac{x}{2\sigma^2}}. \end{aligned} \tag{2.8}$$

In this calculation, we have used the approximation $g|\langle\hat{A}\rangle_w| \ll 1$ which is sometimes called "AAV approximation". The probe position distribution is evaluated as

$$f_w(x|g) = \frac{|\langle x|\psi_w\rangle|^2}{\int dx|\langle x|\psi_w\rangle|^2} \approx \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} e^{-\frac{1}{2\sigma^2}\left(x - g\mathrm{Re}\langle\hat{A}\rangle_w\right)^2}. \tag{2.9}$$

Thus we can obtain the real part of the weak value from the expectation value of the probe position as

$$\langle\hat{x}\rangle_w := \frac{\langle\psi_w|\hat{x}|\psi_w\rangle}{\langle\psi_w|\psi_w\rangle} \approx g\mathrm{Re}\langle\hat{A}\rangle_w. \tag{2.10}$$

Hence we can experimeantly get the weak value for a given interaction strength.

We remark that with those approximations, we can also obtain the imaginary part of the weak value by measuring the final probe wave function in the momentum space given as

$$\begin{aligned}\langle p|\psi_w\rangle \approx & \langle f|i\rangle(1 - ig\langle\hat{A}\rangle_w p)\langle p|\psi\rangle \approx \langle f|i\rangle e^{-ig\langle\hat{A}\rangle_w p}\langle p|\psi\rangle \\ = & \langle f|i\rangle \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{4}} e^{-\sigma^2\left(p - \frac{g}{2\sigma^2}\mathrm{Im}\langle\hat{A}\rangle_w\right)^2 + \frac{g^2}{4\sigma^2}\mathrm{Im}\langle\hat{A}\rangle_w^2 - ig\mathrm{Re}\langle\hat{A}\rangle_w p} \end{aligned} \tag{2.11}$$

which leads to the expectation value of the final probe momentum

$$\langle\hat{p}\rangle_w := \frac{\langle\psi_w|\hat{p}|\psi_w\rangle}{\langle\psi_w|\psi_w\rangle} \approx \frac{g}{2\sigma^2}\mathrm{Im}\langle\hat{A}\rangle_w. \tag{2.12}$$

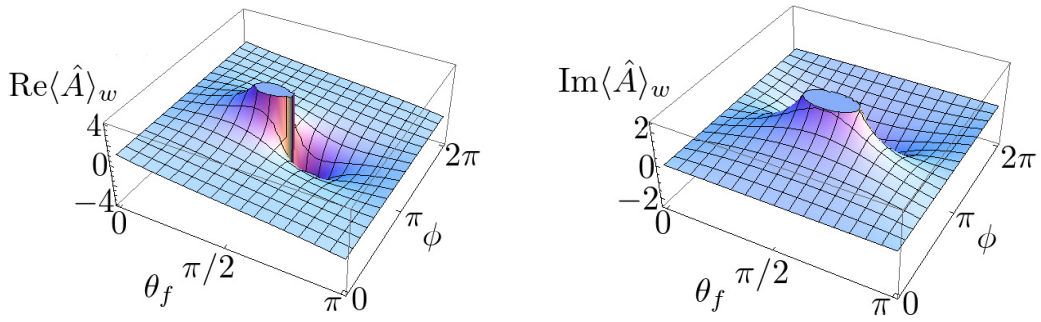


Figure 2.2: Plots of Left: $\mathrm{Re}\langle\hat{A}\rangle_w$ and Right: $\mathrm{Im}\langle\hat{A}\rangle_w$ with the axes θ_f and $\phi(=\phi_f-\phi_i)$. In each plots we fix $\theta_i=\pi/2$.

Let us see the weak value when the measured system, for example, is two-state system. The initial and final states are given as

$$|i\rangle = \cos\frac{\theta_i}{2}|+\rangle + \sin\frac{\theta_i}{2}e^{i\phi_i}|-\rangle, \tag{2.13}$$

$$|f\rangle = \cos\frac{\theta_f}{2}|+\rangle + \sin\frac{\theta_f}{2}e^{i\phi_f}|-\rangle, \tag{2.14}$$

where $0 \le \theta_{i,f} < \pi$ and $0 \le \phi_{i,f} \le 2\pi$, and the observable is $\hat{A} = |+\rangle\langle+| - |-\rangle\langle-|$. The weak value (2.7) is calculated as

$$\langle\hat{A}\rangle_w = \frac{\cos\theta_i + \cos\theta_f + i\sin\theta_i\sin\theta_f\sin(\phi_f-\phi_i)}{1+\cos\theta_i\cos\theta_f+\sin\theta_i\sin\theta_f\cos(\phi_f-\phi_i)}. \tag{2.15}$$

As shown in Fig. 2.2, the weak value can be arbitrarily large by tuning the postselected state almost orthogonal to the preselected state, i.e., $\theta_i+\theta_f \approx \pi, \phi_f-\phi_i \approx \pi$.

We have reviewed the conventional and the weak measurements, the targets of which are the initial state $|i\rangle$ of the measured system and the weak value, respectively. On the other hand, the aim of the WVA is the extraction of the coupling constant g between the measured system and the measuring probe. Under the given weak value, we can obtain the coupling constant from the final probe expectation values (2.10) and (2.12). While the conventional measurement can also provide the interaction strength under the given initial state $|i\rangle$, the "intuitive" merit of the WVA method is that the weak value can make the shift of the expectation value outside the eigenvalues ranges [19]. Therefore we will extract the information of the coupling constant even if it is very small. Measurments of the weak values and the WVA have been done in several experiments [18]. In Chap. 6, we exemplify the experiment using a single birefringent crystal demonstrated by Ritchie *et al.* [20] for the weak measurement and the WVA.

We remark that, as seen in the definition (2.7), the weak value becomes arbitrarily

large for the almost orthogonal pair of the initial and final states. However, the expectation value does not become infinitely large for the arbitrarily large weak value like Eqs. (2.10) and (2.12), because the AAV approximation breaks down. Some researchers have evaluated the expectation values with the full-order caluculation under the particular assumption [32–34]. We see it in the next section.

2.3 Amplified shift in full-Order Calculation

We see the exact case of the weak measurement for an arbitrary interaction strength g and the weak value $\langle\hat{A}\rangle_w$. To calculate the expectation values of the probe position and momentum without any approximation, here we assume that the observable $\hat{A}$ satisfies the property $\hat{A}^2 = 1$, i.e., $\hat{A} = |+\rangle\langle+| - |-\rangle\langle-|$, and the initial probe retains the Gaussian profile (2.1). The final probe state after the postselecion is calculated as

$$\begin{aligned}|\psi_w\rangle &= \langle f|e^{-ig\hat{A}\otimes\hat{p}}|i\rangle|\psi\rangle \\ &= \langle f|[\cos(g\hat{A}\otimes\hat{p}) - i\sin(g\hat{A}\otimes\hat{p})]|i\rangle|\psi\rangle \\ &= \langle f|(\cos g\hat{p} - i\hat{A}\sin g\hat{p})|i\rangle|\psi\rangle \\ &= \langle f|i\rangle(\cos g\hat{p} - i\langle\hat{A}\rangle_w \sin g\hat{p})|\psi\rangle. \end{aligned} \tag{2.16}$$

For later convenience, we introduce $B(\hat{p}) := \cos g\hat{p} - i\langle\hat{A}\rangle_w \sin g\hat{p}$. First, we consider the probability distribution of the final probe momentum $\tilde{f}_w(p|g)$, which is calculated as

$$\tilde{f}_w(p|g) = \frac{|\langle p|\langle f|\hat{U}|i\rangle|\psi\rangle|^2}{|\langle f|\hat{U}|i\rangle|\psi\rangle|^2} = \frac{|\langle p|B(\hat{p})|\psi\rangle|^2}{|B(\hat{p})|\psi\rangle|^2} = \frac{|B(p)\tilde{\psi}(p)|^2}{\int dp|B(p)\tilde{\psi}(p)|^2} \tag{2.17}$$

and

$$|B(p)|^2 = \frac{1}{2}(1 + |\langle\hat{A}\rangle_w|^2) + \frac{1}{2}(1 - |\langle\hat{A}\rangle_w|^2)\cos 2gp + \mathrm{Im}\langle\hat{A}\rangle_w \sin 2gp. \tag{2.18}$$

Here we provide the calculation formula for general real values a and $b(>0)$ as

$$\int dp \cos ap\ e^{-bp^2} = \int dp\ e^{iap}e^{-bp^2} = \int dp\ e^{-b(p+i\frac{a}{2b})^2 - \frac{a^2}{4b}} = \left(\frac{\pi}{b}\right)^{\frac{1}{2}} e^{-\frac{a^2}{4b}}. \tag{2.19}$$

Thus the denominator is calculated as

$$\begin{aligned}\int dp|B(p)\tilde{\psi}(p)|^2 &= \frac{1}{2}(1+|\langle\hat{A}\rangle_w|^2)+\frac{1}{2}(1-|\langle\hat{A}\rangle_w|^2)\left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{2}}\int dp\cos 2gp\ e^{-2\sigma^2p^2}\\ &= \frac{1}{2}(1+|\langle\hat{A}\rangle_w|^2)+\frac{1}{2}(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}} := \frac{\mathcal{Z}}{2}. \end{aligned} \tag{2.20}$$

Thus the probability distribution is

$$\tilde{f}_w(p|g) = \frac{2|B(p)\tilde{\psi}(p)|^2}{\mathcal{Z}}. \tag{2.21}$$

The expectation value of the final probe momentum is calculated as

$$\langle\hat{p}\rangle_w = \int dp\, p\tilde{f}_w(p|g) = \frac{2\int dp\ p|B(p)\tilde{\psi}(p)|^2}{\mathcal{Z}}. \tag{2.22}$$

Because the numerator is

$$\begin{aligned}\int dp\ p|B(p)\tilde{\psi}(p)|^2 &= \mathrm{Im}\langle\hat{A}\rangle_w\left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{2}}\int dp\ p\sin 2gp\ e^{-2\sigma^2p^2}\\ &= \mathrm{Im}\langle\hat{A}\rangle_w\left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{2}}\mathrm{Im}\left[\int dp\ pe^{2igp}\ e^{-2\sigma^2p^2}\right]\\ &= \mathrm{Im}\langle\hat{A}\rangle_w\left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{2}}\mathrm{Im}\left[\int dp\ pe^{-2\sigma^2(p-i\frac{g}{2\sigma^2})^2-\frac{g^2}{2\sigma^2}}\right]\\ &= \frac{g}{2\sigma^2}\mathrm{Im}\langle\hat{A}\rangle_w e^{-\frac{g^2}{2\sigma^2}}, \end{aligned} \tag{2.23}$$

the expectation value of the final probe momentum becomes

$$\langle\hat{p}\rangle_w = 2\frac{g}{2\sigma^2}\frac{\mathrm{Im}\langle\hat{A}\rangle_w e^{-\frac{g^2}{2\sigma^2}}}{\mathcal{Z}} = \frac{g}{\sigma^2}\frac{\mathrm{Im}\langle\hat{A}\rangle_w e^{-\frac{g^2}{2\sigma^2}}}{(1+|\langle\hat{A}\rangle_w|^2)+(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}}. \tag{2.24}$$

Next we evaluate the position distribution $f_w(x|g)$, which can be calculated by the inverse Fourier transform as

$$f_w(x|g) = \frac{2}{\mathcal{Z}}\left|\frac{1}{\sqrt{2\pi}}\int dpB(p)\tilde{\psi}(p)e^{ipx}\right|^2. \tag{2.25}$$

The integration becomes

$$
\begin{aligned}
&\int dpB(p)\tilde{\psi}(p)e^{ipx} \\
&= \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{4}} \int dp(\cos gp - i\langle\hat{A}\rangle_w \sin gp)e^{-\sigma^2p^2}e^{ipx} \\
&= \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{4}} \int dp(\cos gp - i\langle\hat{A}\rangle_w \sin gp)e^{-\sigma^2\left(p-i\frac{x}{2\sigma^2}\right)^2-\frac{x^2}{4\sigma^2}} \\
&= \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{4}} e^{-\frac{x^2}{4\sigma^2}} \int dp\left[\cos\left(gp+i\frac{gx}{2\sigma^2}\right) - i\langle\hat{A}\rangle_w \sin\left(gp+i\frac{gx}{2\sigma^2}\right)\right] e^{-\sigma^2p^2} \\
&= \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{4}} e^{-\frac{x^2}{4\sigma^2}} \left[\cos i\frac{gx}{2\sigma^2} - i\langle\hat{A}\rangle_w \sin i\frac{gx}{2\sigma^2}\right] \int dp\cos gp\ e^{-\sigma^2p^2} \\
&= \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{4}} e^{-\frac{x^2}{4\sigma^2}} \left[\cosh\frac{gx}{2\sigma^2} + \langle\hat{A}\rangle_w \sinh\frac{gx}{2\sigma^2}\right] \left(\frac{\pi}{\sigma^2}\right)^{\frac{1}{2}} e^{-\frac{g^2}{4\sigma^2}} \\
&= \frac{1}{2}\left(\frac{2\pi}{\sigma^2}\right)^{\frac{1}{4}} [(1+\langle\hat{A}\rangle_w)e^{-\frac{(x-g)^2}{4\sigma^2}} + (1-\langle\hat{A}\rangle_w)e^{-\frac{(x+g)^2}{4\sigma^2}}]. \qquad (2.26)
\end{aligned}
$$

Thus, the position distribution $f_w(x|g)$ is

$$
\begin{aligned}
f_w(x|g) =& \frac{2}{\sqrt{2\pi\sigma^2}\mathcal{Z}}\left|(1+\langle\hat{A}\rangle_w)e^{-\frac{(x-g)^2}{4\sigma^2}} + (1-\langle\hat{A}\rangle_w)e^{-\frac{(x+g)^2}{4\sigma^2}}\right|^2 \\
=& \frac{2}{\sqrt{2\pi\sigma^2}\mathcal{Z}}\left[(1+|\langle\hat{A}\rangle_w|^2 + 2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x-g)^2}{2\sigma^2}}\right. \\
&\left. +(1+|\langle\hat{A}\rangle_w|^2 - 2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x+g)^2}{2\sigma^2}} + 2(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{x^2+g^2}{2\sigma^2}}\right]. \qquad (2.27)
\end{aligned}
$$

Hence, the expectation value of the final probe position is

$$
\langle\hat{x}\rangle_w = \int dx\ xf_w(x|g) = g\frac{2\mathrm{Re}\langle\hat{A}\rangle_w}{\mathcal{Z}} = g\frac{2\mathrm{Re}\langle\hat{A}\rangle_w}{(1+|\langle\hat{A}\rangle_w|^2) + (1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}}. \qquad (2.28)
$$

We can extract the weak value from the expectations values (2.24) and (2.28). However, we can easily see that the expectations values cannot be infinitely large even if the weak value can be arbitrarily large, because the denominators of Eqs. (2.24) and (2.28) have the second order term of the weak value. This term comes from the higher-order terms of the coupling constant, which is understood as a back action of the measurement. Actually,

we can show the weak value which gives the maximal expectation value as shown below;

$$\frac{\partial\langle\hat{p}\rangle}{\partial\mathrm{Re}\langle\hat{A}\rangle_w} = -g\sigma^2 e^{-\frac{g^2}{2\sigma^2}} \frac{2(1-e^{-\frac{g^2}{2\sigma^2}})\mathrm{Re}\langle\hat{A}\rangle_w \mathrm{Im}\langle\hat{A}\rangle_w}{[(1+|\langle\hat{A}\rangle_w|^2)+(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}]^2}, \tag{2.29}$$

$$\frac{\partial\langle\hat{p}\rangle}{\partial\mathrm{Im}\langle\hat{A}\rangle_w} = 2g\frac{1+e^{-\frac{g^2}{2\sigma^2}}+(1-e^{-\frac{g^2}{2\sigma^2}})(\mathrm{Re}\langle\hat{A}\rangle_w^2-\mathrm{Im}\langle\hat{A}\rangle_w^2)}{[(1+|\langle\hat{A}\rangle_w|^2)+(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}]^2}, \tag{2.30}$$

$$\frac{\partial\langle\hat{x}\rangle_w}{\partial\mathrm{Re}\langle\hat{A}\rangle_w} = 2g\frac{1+e^{-\frac{g^2}{2\sigma^2}}-(1-e^{-\frac{g^2}{2\sigma^2}})(\mathrm{Re}\langle\hat{A}\rangle_w^2-\mathrm{Im}\langle\hat{A}\rangle_w^2)}{[(1+|\langle\hat{A}\rangle_w|^2)+(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}]^2}, \tag{2.31}$$

$$\frac{\partial\langle\hat{x}\rangle_w}{\partial\mathrm{Im}\langle\hat{A}\rangle_w} = \frac{-4g(1-e^{-\frac{g^2}{2\sigma^2}})\mathrm{Re}\langle\hat{A}\rangle_w \mathrm{Im}\langle\hat{A}\rangle_w}{[(1+|\langle\hat{A}\rangle_w|^2)+(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}]^2}. \tag{2.32}$$

We can obtain each maximal shift as

$$\mathrm{Max}\langle\hat{p}\rangle_w = \frac{g}{2\sigma^2}\frac{e^{-\frac{g^2}{2\sigma^2}}}{\sqrt{1-e^{-\frac{g^2}{\sigma^2}}}} \quad \text{for } \mathrm{Re}\langle\hat{A}\rangle_w = 0,\ \mathrm{Im}\langle\hat{A}\rangle_w = \sqrt{\frac{1+e^{-\frac{g^2}{2\sigma^2}}}{1-e^{-\frac{g^2}{2\sigma^2}}}}, \tag{2.33}$$

$$\mathrm{Max}\langle\hat{x}\rangle_w = \frac{g}{\sqrt{1-e^{-\frac{g^2}{\sigma^2}}}} \quad \text{for } \mathrm{Re}\langle\hat{A}\rangle_w = \sqrt{\frac{1+e^{-\frac{g^2}{2\sigma^2}}}{1-e^{-\frac{g^2}{2\sigma^2}}}},\ \mathrm{Im}\langle\hat{A}\rangle_w = 0. \tag{2.34}$$

For an appropriate weak value, $\mathrm{Max}\langle\hat{x}\rangle_w$ can be arbitrarily large with the large σ, while $\mathrm{Max}\langle\hat{p}\rangle_w$ has a local maximal value at σ satisfying $\sqrt{1-e^{-g^2/\sigma^2}} = g^2/2\sigma^2$.

We note that the shift of the expectation value depends on the initial probe wave function. Here we present one of examples of the position expectation value as

$$\langle\hat{x}\rangle_w = g\frac{2\mathrm{Re}\langle\hat{A}\rangle_w}{(1+|\langle\hat{A}\rangle_w|^2)+(1-|\langle\hat{A}\rangle_w|^2)\int dp\cos 2gp|\tilde{\psi}(p)|^2}, \tag{2.35}$$

which is derived in the case that the initial probe wave function $\tilde{\psi}(p)$ is an even real function and converges to 0 at $x \to \pm\infty$ including the Gaussian probe case. We show the derivation in Appendix A.

Let us see the figures of the position and momentum distribution and its expectation values given by the weak measurement. For an illustration, we consider the case that the measured system is the two-state ($|\pm\rangle$) system, the observable of which is $\hat{A} = |+\rangle\langle+| - |-\rangle\langle-|$. The initial and final states of the measured system and the weak value are given as Eqs. (2.13), (2.13), and (2.15). From Eqs. (2.21), (2.24), (2.27), and (2.28)

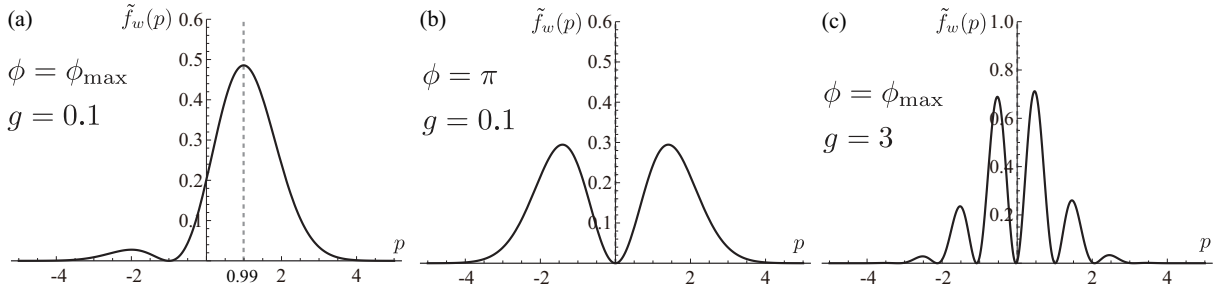


Figure 2.3: The plots of the final probe distribution in the momentum space (2.37) and its expectation values (2.38) in the three cases. The values of $\phi(:= \phi_f - \phi_i)$ and g are different in each cases. $\phi_{\max}(\approx 2.94)$ in (a) means ϕ which maximizes the expectation value. We fix the other parameters as $\theta_i = \theta_f = \pi/2$, and $\sigma = 0.5$.

and using

$$\mathcal{Z} = 2\frac{1 + \cos\theta_i \cos\theta_f + \sin\theta_i \sin\theta_f \cos(\phi_f - \phi_i) e^{-\frac{g^2}{2\sigma^2}}}{1 + \cos\theta_i \cos\theta_f + \sin\theta_i \sin\theta_f \cos(\phi_f - \phi_i)}, \tag{2.36}$$

we obtain the distributions and the expectation values in another form as

$$\begin{aligned}
\mathcal{Z}_N &:= 1 + \cos\theta_i \cos\theta_f + \sin\theta_i \sin\theta_f \cos(\phi_f - \phi_i) e^{-\frac{g^2}{2\sigma^2}}, \\
\tilde{f}_w(p|g) &= \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{2}} \frac{e^{-2\sigma^2 p^2}}{\mathcal{Z}_N} [1 + \cos\theta_i \cos\theta_f + \sin\theta_i \sin\theta_f \cos(2gp - \phi_f + \phi_i)], \qquad (2.37) \\
\langle \hat{p} \rangle_w &= \frac{g}{2\sigma^2 \mathcal{Z}_N} [\sin\theta_i \sin\theta_f \sin(\phi_f - \phi_i) e^{-\frac{g^2}{2\sigma^2}}], \qquad (2.38) \\
f_w(x|g) &= f_{w1}(x|g) + f_{w2}(x|g) + f_{w3}(x|g), \qquad (2.39) \\
f_{w1}(x|g) &:= \frac{1}{\sqrt{2\pi\sigma^2}\mathcal{Z}_N} (2\cos^2\frac{\theta_i}{2} \cos^2\frac{\theta_f}{2} e^{-\frac{(x-g)^2}{2\sigma^2}}), \\
f_{w2}(x|g) &:= \frac{1}{\sqrt{2\pi\sigma^2}\mathcal{Z}_N} (2\sin^2\frac{\theta_i}{2} \sin^2\frac{\theta_f}{2} e^{-\frac{(x+g)^2}{2\sigma^2}}), \\
f_{w3}(x|g) &:= \frac{1}{\sqrt{2\pi\sigma^2}\mathcal{Z}_N} [\sin\theta_i \sin\theta_f \cos(\phi_f - \phi_i) e^{-\frac{x^2+g^2}{2\sigma^2}}], \\
\langle \hat{x} \rangle_w &= \frac{g}{\mathcal{Z}_N} (\cos\theta_i + \cos\theta_f), \qquad (2.40)
\end{aligned}$$

respectively.

In Fig. 2.3, we plot the final probe distribution in the momentum space (2.37) and its expectation values (2.38) in the three cases. (a) shows the case that the expectation value is maximally shifted in the specific condition for setting the phase $\phi_f - \phi_i$ which meets Eq. (2.33). Because the expectation value becomes $\langle \hat{p} \rangle_w \sim 0.99$, we can see that the coupling strength $g = 0.1$ is amplified. The weak value in (a) is computed as $\langle \hat{A} \rangle_w \sim 10i$. On the

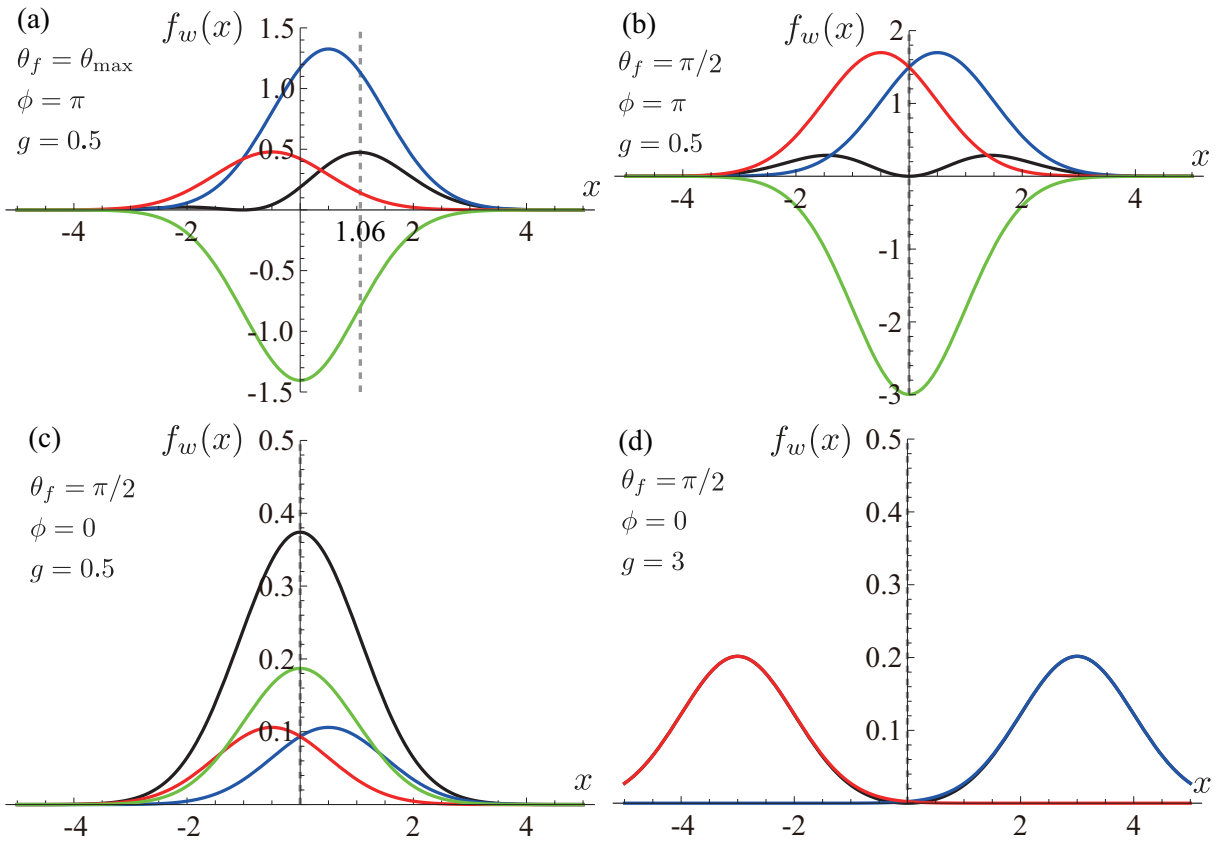


Figure 2.4: The plots of the final probe position distribution (2.39) and the expectation value of which (2.40). We depict the four curves, $f_w(x|g)$ in black, $f_{w1}(x|g)$ in blue, $f_{w2}(x|g)$ in red, and $f_{w3}(x|g)$ in green. The dashed vertical lines indicate the expectation value. The parameters of θ_f, ϕ, and g are tuned in each cases. $\theta_{\max}(\approx 1.08)$ in (a) means θ_f maximizing the expectation value. We have set $\theta_i = \pi/2, \sigma = 1$ in all cases.

other hand, in (b), we set the phase which gives the two-equal-height-peak distribution and the indeterminate weak value. In this case, the peak momentums are $p \sim \pm 1.5$ larger than the expectation value in (a), while one in (b) is zero. We set the strong interaction in (c), which shows the multimodal function. There are also the peaks, the momentums of which are large, but it is somewhat difficult to observe due to their probabilities smaller than the one nearby zero-momentum.

Figure 2.4 displays the final probe position distributions (2.39) and the expectation values of which (2.40) in the four cases. The case (a) gives the weak value $\langle \hat{A} \rangle_w \approx 4$ and the maximal shift to the expectation value of position, the value of which is about 1.06 while the interaction strength is 0.5. Hence, we can see the amplification effect. On the other hand, in the case (b), the distribution $f_w(x|g)$ has two peaks at $x \approx \pm 2$ with the zero expectation value. The weak value becomes infinity for the postselected state completely orthogonal to the preselected state, which breaks the AAV approximation

$g|\langle\hat{A}\rangle_w| \ll 1$. The authors of Ref. [19] said that breaking the AAV approximation brings the notable amplification effect. The third term of the distribution $f_{w3}(x|g)$ contributes to the amplification by its negative canceling the overlapped region of $f_{w1}(x|g)$ and $f_{w2}(x|g)$. The same holds in the case (a). Actually, the distribution $f_w(x|g)$ in (c) does not exhibit the amplification effect due to the positive $f_{w3}(x|g)$. We can find that $f_{w3}(x|g)$ contributes to the amplification. (d) represents the strong coupling constant case, which brings $f_{w3}(x|g) = 0$ regardless of the weak value. So, the strong coupling constant cancel the amplification effect.

In this section, we see that the amplification by the weak measurement has an upper bound due to the higher-order terms in the coupling constant. In other words, the back action of the measurement is significant, while we can amplify the output larger than the one obtained by the conventional measurement. In next section, we see the WVA from the viewpoint of the parameter estimation.

2.4 Parameter Estimation

The purpose of the WVA in a broad sense is an extraction of the information about the interaction strength by the amplification. In the statistical inference, there is a prescription for estimating the parameter provided by the estimation theory. Some researchers applied the estimation theory to evaluate the capability of the WVA for estimating the parameter in each different situation [40–44]. Meanwhile, there is a criticism to the WVA that the large amplification makes a number of obtained data small. The survival rate of the data is calculated as the transition probability:

$$P = |\langle f|e^{-ig\hat{A}\otimes\hat{p}}|i\rangle|\psi\rangle|^2 = \frac{\mathcal{Z}}{2}|\langle f|i\rangle|^2 \tag{2.41}$$

which is sometimes called as the success probability of the postselection. By the postselection for a large amplification, the probability becomes small. Some researchers argue that this point is the disadvantage of the WVA in respect to estimating the coupling constant. Especially, the authors of Ref. [41] showed the essence by evaluating the quantum Fisher information that the postselection cannot increase the information due to the small success probability of the postselection. The Fisher information gives the lower bound of the mean square error of an estimator. Namely, the WVA is worse for estimating the parameter than the conventional measurement.

In this section, we see the Cramér-Rao inequality which gives the relation between the mean square error and the Fisher information. We compare the Fisher information of the probability distributions obtained by that of the conventional measurement and the weak measurement. Here we consider the classical Fisher information instead of the quantum one. When we evaluate the quantum Fisher information, we assume that we take the measurement which maximizes the Fisher information. However, to see the essence, it is enough to calculate the classical Fisher information, which is calculated from the obtained probability distribution $f(x)$. As stated in Ref. [41], we need to get the classical Fisher information for a practical comparison of the estimation capabilities of the weak measurement and the conventional measurement.

Firstly, we introduce an unbiased estimator $\tilde{\theta}(x)$ satisfying

$$\langle \tilde{\theta}(x) \rangle = \theta \text{ for all } \theta, \tag{2.42}$$

where θ is a true value of the parameter that we want to measure. When the estimator takes a close value of the true value on average, we can say that the estimation works well. We employ the mean square error $\langle (\tilde{\theta}(x) - \theta)^2 \rangle$ as an indicator of the estimation. The Cramér-Rao inequality tells us the lower bound of the mean square error as

$$\langle (\tilde{\theta}(x) - \theta)^2 \rangle \geq \frac{1}{nI(\theta)}, \tag{2.43}$$

where $I(\theta)$ is the classical Fisher information defined as

$$I(\theta) = \int dx \, [\partial_\theta \log f(x|\theta)]^2 f(x|\theta) \tag{2.44}$$

and n is the number of the data, which are obtained from independent identical distributions. We give the proof of the Cramér-Rao inequality in Appendix B.1. Therefore, the Fisher information can be an indicator of the estimation capability of the measurement, although it is not always equal to the mean square error.

Here we give the Fisher information for each measurement in the case that we can analytically calculate. Like the previous sections, we take the two-state ($|\pm\rangle$) as the measured system, the observable $\hat{A} = |+\rangle\langle+| - |-\rangle\langle-|$, and the Gaussian probe (2.1). We set the estimation problem that we estimate the value of the coupling constant from the position distribution obtained by measurements. For the conventional measurement, we assume that the initial state of the measured system is $|i\rangle = |+\rangle$. The final probe

distribution (2.5) becomes

$$f^c(x|g) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-g)^2}{2\sigma^2}}, \tag{2.45}$$

which gives the Fisher information

$$I^c(g) = \frac{1}{\sigma^2}. \tag{2.46}$$

Next, we look at the weak measurement. We assume that $\theta_i = \theta_f = \pi/2$ and $\cos(\phi_f - \phi_i) = \epsilon$. Moreover, we consider only the case $\epsilon = \pm 1$ for an analytical calculation, where $\epsilon = 1$ and $\epsilon = -1$ correspond with (c) and (b) in Fig. (2.4), respectively. From Eq. (2.39) with these assumptions, the final probe distribution becomes

$$f^w_\epsilon(x|g) = \frac{1}{2\sqrt{2\pi\sigma^2}} \frac{e^{-\frac{(x-g)^2}{2\sigma^2}} + e^{-\frac{(x+g)^2}{2\sigma^2}} + 2\epsilon e^{-\frac{x^2+g^2}{2\sigma^2}}}{1 + \epsilon e^{-\frac{g^2}{2\sigma^2}}}, \tag{2.47}$$

which gives the Fisher information[1]

$$I^w_\epsilon(g) = \frac{1}{\sigma^2} \frac{1 + \epsilon \frac{g^2}{\sigma^2} e^{-\frac{g^2}{2\sigma^2}} - e^{-\frac{g^2}{\sigma^2}}}{(1 + \epsilon e^{-\frac{g^2}{2\sigma^2}})^2}. \tag{2.48}$$

We describe these calculations in Appendix B.2.

Here we pay attention to the right-hand side of the Cramér-Rao inequality (2.43) which has the number of obtained date n. In some case, we should consider the success probability of the postselection (2.41), which would make the number of the detected data small. If the number in the conventional measurement is n, the one in the weak measurement becomes nP, where

$$P = \frac{1 + \epsilon e^{-\frac{g^2}{2\sigma^2}}}{2}. \tag{2.49}$$

Figure 2.5 shows the plots of the Fisher information I^c and I^w_ϵ. (a) shows that the Fisher information given in the weak measurement with $\epsilon = 1$ can exceed the one in the conventional measurement ($1/\sigma^2$). However, in (b) which includes the factor of the success probability of the postselection, we can find the inequality with respect to the

[1] Although the authors of Ref. [41] forget the factor of ϵ in the denominator of Eq. (2.48), their conclusion remains true.

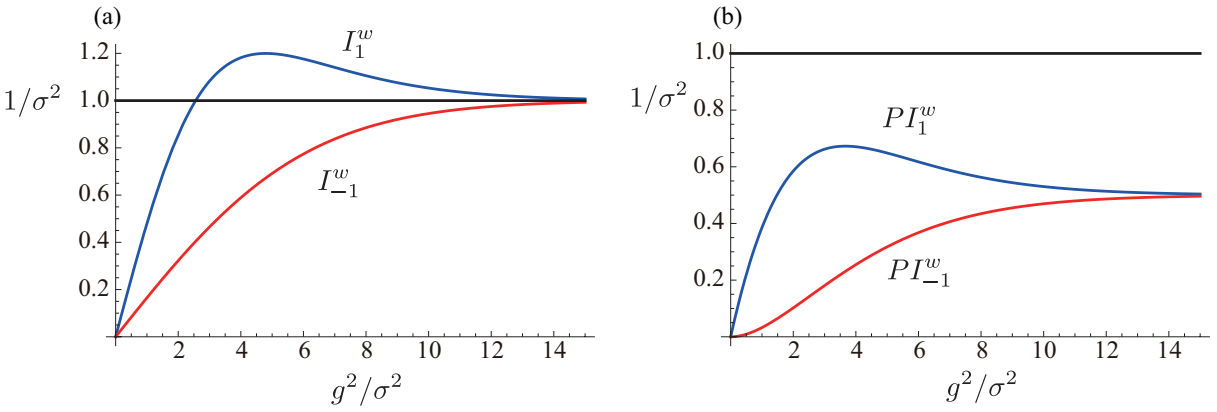


Figure 2.5: Plots of the Fisher information of the conventional measurement (black line) and the weak measurement (blue curve for $\epsilon = 1$ and red curve for $\epsilon = -1$). The vertical axis is normalized by $1/\sigma^2$ and the horizontal axis indicates g^2/σ^2. (a) shows a simple comparison of the Fisher information, and (b) includes the factor of the success probability of the postselection.

Fisher information as

$$I^c > PI_\epsilon^w, \tag{2.50}$$

regardless of the value of ϵ. We remark again that the Fisher information only represents the lower bound of the mean square error of the estimator. Here we put the Cramér-Rao inequalities in the conventional measurement and the weak measurement as

$$\langle(\tilde{g}_c(x) - g)^2\rangle_c \geq \frac{1}{nI^c(g)}, \tag{2.51}$$

$$\langle(\tilde{g}_w(x) - g)^2\rangle_w \geq \frac{1}{PnI_\epsilon^w(g)}, \tag{2.52}$$

where $\tilde{g}_c(x)$ and $\tilde{g}_w(x)$ are the estimators for each measurement. If we want to compare the mean square errors, we need to show that the Cramér-Rao inequality (2.43) has equality. When the probability distribution is Gaussian and we want to estimate the mean value of the distribution, we have the uniformly minimum variance unbiased (UMVU) estimator, which gives the equality of the Cramér-Rao inequality. In the Gaussian distribution, the variance of the sample average $\bar{X}$ is equal to the population variance σ divided by sample

number n^2. Namely,

$$\langle(\bar{X}-g)^2\rangle_c = \frac{\sigma^2}{n} = \frac{1}{nI^c(\theta)}, \tag{2.53}$$

which imply that the sample average $\bar{X}$ is the UMVU estimator. Hence the Cramér-Rao inequality (2.43) has equality in the conventional measurement case, when $\tilde{g}_c(x) = \bar{X}$. From this fact and Eqs. (2.50), (2.51), and (2.52), the inequality

$$\langle(\tilde{g}_c(x)-g)^2\rangle_c < \langle(\tilde{g}_w(x)-g)^2\rangle_w \tag{2.54}$$

holds for $\tilde{g}_c(x) = \bar{X}$, which indicates that the weak measurement is worse for estimating the coupling constant than the conventional measurement when the number of trials in an experiment are same in the two measurement. Although we have shown Eq. (2.54) for the Gaussian probe, it is not guaranteed that Eq. (2.54) holds for other probe cases, because an unbiased estimator such as minimizing the mean square error is not always present[3]. We remark that the derived inequality (2.54) holds even for the finite n[4]. Note that we also evaluate the Fisher information with the probe distribution in the momentum space. We give the result in Appendix B.2. In Ref. [43], they compared the Fisher information of the final probe distribution in the position space given by the conventional measurement with one of the final probe distribution in the momentum space given by the weak measurement. They concluded that the former is better than the latter.

Here we focus on the Fishier information given by the weak measurement. We show the two cases, the case of $\epsilon = 1$ which gives no amplification as seen in Fig. 2.4 (c) and the

[2] $\langle(\bar{X}-g)^2\rangle = \mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{\sum_i^n X_n}{n}\right) = \frac{1}{n^2}\mathrm{Var}\left(\sum_i^n X_n\right) = \frac{1}{n^2}\sum_i^n \mathrm{Var}(X_n) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$, where X_n is a sample value.

[3] When there is not a UMVU estimator, the maximum likelihood estimator (MLE) is usually used for the practical reason. According to the asymptotic theory, roughly speaking, by using the MLE with the infinitely large number of data, the mean square error approaches to the inverse of the Fisher information as stochastic convergence. This property is referred as the asymptotic efficiency of the maximum likelihood estimator. Hence, if we have the infinitely large number of data, the comparison of the Fisher information is meaningful for as evaluating the estimation capability, even if a UMVU estimator does not exst. On the other hand, this conclusion will be meaningless as the comparison of the estimation capabilities when the number of data is finite, although the inequality (2.50) holds.

[4] The authors of Ref. [43] shows the inequality (2.50) for a probe spatial wave function which is real valued under the assumption that the weak coupling constant ($g \ll 1$) and the AAV approximation ($g|\langle\hat{A}\rangle_w| \ll 1$), and the asymptotic efficiency ($n \to \infty$). In addition, they give the inequality (2.50) for the Gaussian probe without any approximation, while using the asymptotic efficiency. The authors of Ref. [42] have shown it in Supplemental Material that the quantum Fisher information given by the weak measurement is smaller than one given by the conventional measurement not only in the asymptotic case but also the finite data case based on the Chernoff bound. There is no mention of the equality of the Cramér-Rao inequality.

case $\epsilon = -1$ which provides a large amplification as seen in Fig. 2.4 (b). As seen in Fig. 2.5, the Fishier information in $\epsilon = -1$ cannot overcome the one in $\epsilon = 1$ whether or not we mind the success probability of the postselection. Although we have to care the equality of the Cramér-Rao inequality to discuss the estimation capability as stated in above, we can regard that the amplification effect primarily has a disadvantage for estimating the coupling constant regardless of the success probability of the postselection.

2.5 Summary of this chapter

In this chapter, we have reviewed the concept of the weak measurement and the weak-value amplification originally proposed by Aharonov, Albert and Vaidman [17]. In a proposal in Ref. [17], the authors considered only the case that the coupling constant is weak. As we have seen in Sec. 2.2, by the weak measurement, i.e., the indirect quantum measurement with postselecting the final state of the measured system, the weak value appears in the shift of the expectation value of the measuring probe. The weak value can be complex and arbitrarily large by tuning pre- and postselected state of the measured system. In the WVA, we apply this property for amplifying the output and measuring the coupling constant between the measured system and the measuring probe [19]. Hence some researchers are developing the WVA for a precise measurement technique [14–16]. Actually, some researchers demonstrated the WVA by several experiments [18]. Note that calculating the probability distribution without a weak coupling approximation in a certain situation, we can analytically show that the amplification has an upper bound for the higher-order terms in the coupling constant [32–34]. Simultaneously, the output of the weak measurement depends on the initial probe wave function. It means that we will engineer the probe wave function providing a large amplification and a small variance for a certain experiment. If we manufacture such a probe wave function, the application of the WVA may be expanded. We discuss this topic in the next chapter.

Meanwhile, as we have mentioned, some researchers argue that the postselection makes the number of detectable data small, which brings the disadvantage for the WVA. To see this issue, in this thesis, we have exemplified the estimation capability of the weak measurement evaluated by the classical Fisher information, which is also disclosed in Ref. [41]. We have shown that the postselection reduces the classical Fisher information.

Furthermore, in the case of the parameter estimation, it seems that the amplified shift is not utilized even if the success probability of the postselection is neglected. Thus

we consider another statistical inference problem, in which we can intuitively find the merit of the amplification. We pick up the statistical hypothesis testing method, which is usually used for analyzing the signal detection in a physical experiment [53–55]. We review the standard hypothesis testing in Chap. 4. Also we establish the testing method in the weak measurement and compare the testing capability of the two measurements in Chaps. 5 and 6.

In the subsequent chapters, we discuss these two topics, the probe engineering and the hypothesis testing to develop the WVA technique for a precising measurement. Here we remark that when we discuss practical accuracy of the measurement, we need to consider a technical noise. Not mentioning about a technical noise in the main body, we introduce the photon shot noise which is one of the fundamental noises in an optical interferometer in Appendix C.

Chapter 3

Optimal Probe Wave Function in Weak-Value Amplification

3.1 Preface

In this chapter, we analytically derive the optimal probe wave function for the fixed experimental setup, i.e., the given coupling constant and tuned the pre- and postselected states that we can calculate the weak value before the experiment. The wave function gives not only the infinitely large expectation value but also the zero variance of the final probe position in principle. We have the two ways to derive the optimal probe wave function. One is that we derive the function which gives the maximum expectation value of the final probe position. The other is that we derive the function which gives the minimum variance of the final probe position. In both ways, we use the Lagrange multiplier method. In this optimization, we assume that the observable $\hat{A}$ satisfies $\hat{A}^2 = 1$ and $\mathrm{Re}\langle\hat{A}\rangle_w \neq 0$. This chapter is based on the research [58, 59].

3.2 Property of optimal probe

In this section, we show the optimal probe wave function and consider its properties and connotations preparatory for derivation. The optimal initial probe wave function in the

momentum space[1] is given as

$$\tilde{\xi}(p) = \langle p|\xi\rangle = \sqrt{\frac{g|\mathrm{Re}\langle\hat{A}\rangle_w|}{\pi}}\frac{\exp\left[-i\frac{g\left(|\langle\hat{A}\rangle_w|^2+1\right)}{2\mathrm{Re}\langle\hat{A}\rangle_w}p\right]}{\cos gp - i\langle\hat{A}\rangle_w \sin gp} = \sqrt{\frac{g|\mathrm{Re}\langle\hat{A}\rangle_w|}{\pi}}B^{-1}(p)\exp\left[-i\langle\hat{x}\rangle_\xi p\right], \tag{3.1}$$

the support of which is $-\pi/2g \leq p \leq \pi/2g$. The function gives the position expectation value of the initial probe $\langle\hat{x}\rangle_\xi = 0$ and the one of the final probe as

$$\langle\hat{x}\rangle_{\xi_w} := g\frac{|\langle\hat{A}\rangle_w|^2+1}{2\mathrm{Re}\langle\hat{A}\rangle_w}, \tag{3.2}$$

which can be arbitrarily large as the weak value becomes large. Since $B^{-1}(p)$ in the optimal probe wave function cancel out the higher-order terms given by the unitary operator, the shift of the expectation value $\langle\hat{x}\rangle_{\xi_w} - \langle\hat{x}\rangle_\xi$ has no upper-bound in contrast with the case of using the Gaussian probe. This is the characteristic feature of the optimal probe form.

According to Eqs. (2.16) and (3.1), the probe wave function after the postselection is

$$\tilde{\xi}_w(p) := \frac{\langle x|\xi_w\rangle}{\sqrt{\langle\xi_w|\xi_w\rangle}} = \sqrt{\frac{g}{\pi}}\exp\left[-i\langle\hat{x}\rangle_{\xi_w}p\right], \tag{3.3}$$

the support of which is also $-\pi/2g \leq p \leq \pi/2g$. Furthermore, we can obtain the final probe wave function in the position space with the inverse Fourier transform as

$$\xi_w(x) = \frac{2g}{\pi}\frac{\sin[\frac{\pi}{2g}(x - \langle\hat{x}\rangle_{\xi_w})]}{x - \langle\hat{x}\rangle_{\xi_w}}. \tag{3.4}$$

[1]We can also derive the function in the position space with the inverse Fourier transform. When $\langle\hat{A}\rangle_w = -1$, the function is given by

$$\xi(x) = \langle x|\xi\rangle = \sqrt{\frac{2g|\mathrm{Re}\langle\hat{A}\rangle_w|}{\pi^2}}\frac{\sin\left[\frac{\pi}{2g}(x - \langle\hat{x}\rangle_{\xi_w} + g)\right]}{x - \langle\hat{x}\rangle_{\xi_w} + g}.$$

When $\langle\hat{A}\rangle_w \neq -1$, the function becomes

$$\xi(x) = \sqrt{\frac{2|\mathrm{Re}\langle\hat{A}\rangle_w|}{\pi^2 g}}\frac{\sin\left[\frac{\pi}{2g}(x - \langle\hat{x}\rangle_{\xi_w} + g)\right]}{1 + \langle\hat{A}\rangle_w}\Psi\left[\frac{1 - \langle\hat{A}\rangle_w}{1 + \langle\hat{A}\rangle_w}, 1, \frac{x - \langle\hat{x}\rangle_{\xi_w} + g}{2g}\right],$$

where $\Psi[z, s, a] := \sum_{n=0}^{\infty}\frac{z^n}{(a+n)^s}$ called Hurwitz-Lerch zeta function. It is noted that the author partially used Mathematica to derive this.

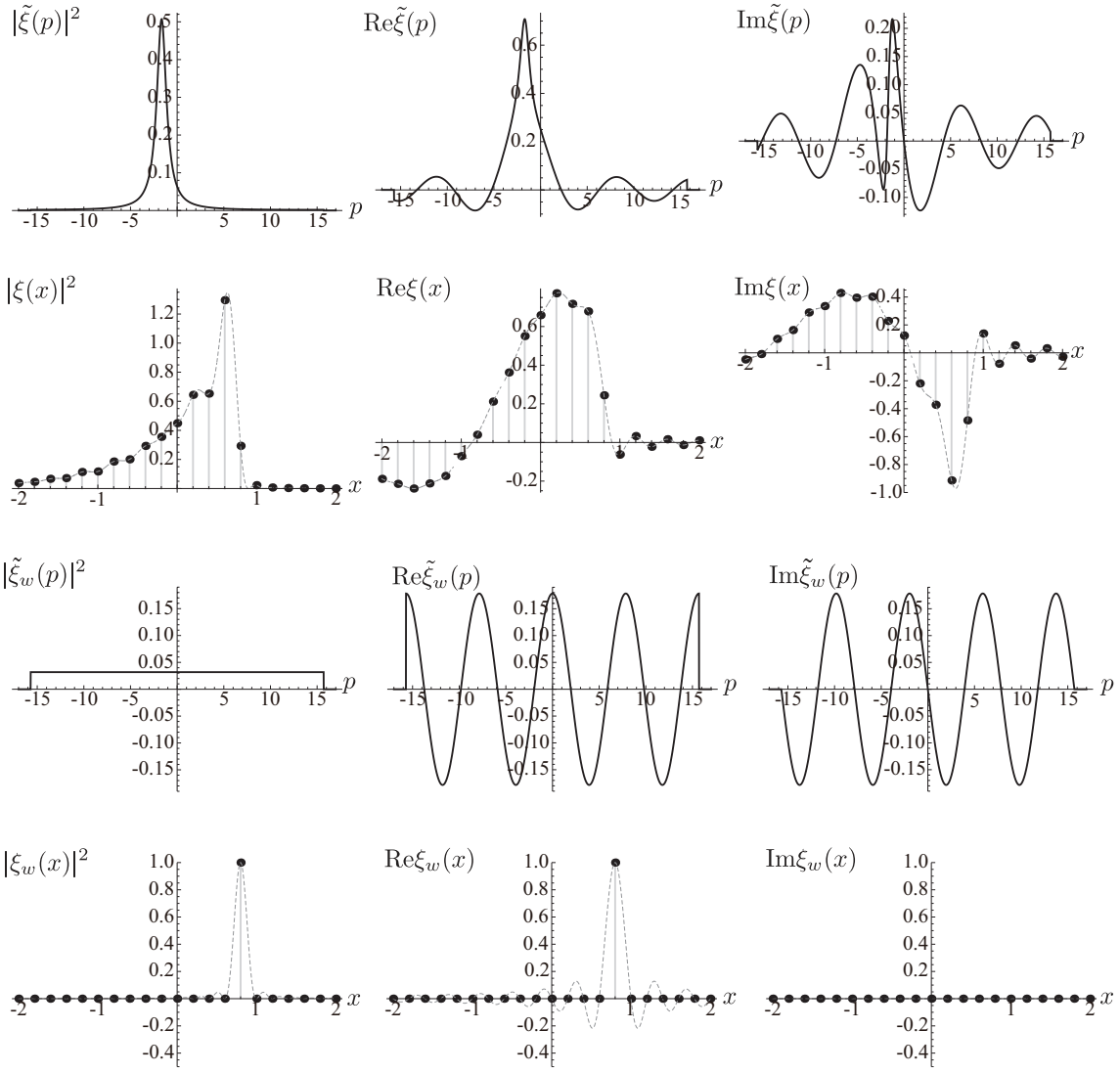


Figure 3.1: The plots of the initial and final optimal probe wave function in the momentum space [$\tilde{\xi}(p)$ and $\tilde{\xi}_w(p)$] and the position space [$\xi(x)$ and $\xi_w(x)$] are shown. The parameters are fixed $g = 0.1$ and $\langle \hat{A} \rangle_w = 2 + 3\sqrt{3}i$. It is noted that the probe wave function in the position space is discrete as dots indicates.

We note that the position x takes discrete values $x = 2gj$ ($j \in \mathbb{N}$) for the finite boundary condition $-\pi/2g \leq p \leq \pi/2g$ of the probe wave function. Therefore, if the final expectation value takes $\langle \hat{x} \rangle_{\xi_w} = 2gk$ ($k \in \mathbb{N}$) by controlling the weak value, the final probe wave function in the position space with the normalization becomes the Kronecker delta as

$$\xi_w(x = 2gj) = \frac{2g}{\pi} \frac{\sin[\frac{\pi}{2g}(2gj - 2gk)]}{2gj - 2gk} = \frac{\sin[(j-k)\pi]}{(j-k)\pi} = \delta_{jk}. \tag{3.5}$$

Hence, the variance of the final probe can be zero.

We remark that the expectation value $\langle\hat{x}\rangle_\xi$ has the lower bound given as

$$|\langle\hat{x}\rangle_\xi| = \frac{|g|}{2}\left(|\mathrm{Re}\langle\hat{A}\rangle_w| + \frac{|\mathrm{Im}\langle\hat{A}\rangle_w|^2+1}{|\mathrm{Re}\langle\hat{A}\rangle_w|}\right) \geq |g|\sqrt{|\mathrm{Im}\langle\hat{A}\rangle_w|^2+1} \geq |g|. \tag{3.6}$$

The minimum $|\langle\hat{x}\rangle_\xi| = |g|$ is given when $\mathrm{Re}\langle\hat{A}\rangle_w = \pm 1$ and $\mathrm{Im}\langle\hat{A}\rangle_w = 0$, which means that the postselected state identical to the one of the eigenstates of the observable $\hat{A}$. In this certain case, we can almost regard the weak measurement as the projective measurement of the system, and the unitary operator by the interaction becomes $e^{\mp ig\hat{p}}$ that gives the shift operator by $\mp g$ to the probe position. It indicates that the weak measurement with the optimal probe wave function always amplifies the shift of the position expectation value more than the projective measurement.

From the expectation value (3.2), we can obtain the convergence

$$\langle\hat{x}\rangle_{\xi_w}|\langle f|i\rangle| \to \frac{g}{2} \tag{3.7}$$

as the postselected state $|f\rangle$ approaches the state orthogonal to the preselected state $|i\rangle$. From this relation, we can see that it is possible to obtain the coupling constant g by extrapolation.

3.3 Derivation

We have two ways to derive the optimal probe wave function (3.1) by using the Lagrange multiplier method. One is by maximizing the shift of the probe position expectation value, and the other is by minimizing the position variance of the final probe distribution.

3.3.1 Maximization of the shift of the expectation value of the probe position

Here, we consider the probe wave function $\tilde{\xi}(p)$ which maximize the shift of the expectation value of the probe position $\langle\hat{x}\rangle_{\xi_w} - \langle\hat{x}\rangle_\xi$. We set the following Lagrangian which gives the extremal value of the expectation value of the final probe position $\langle\hat{x}\rangle_{\xi_w}$:

$$\mathcal{L}[\tilde{\xi}(p), \tilde{\xi}^*(p), \lambda_\mathcal{L}] := \langle\hat{x}\rangle_{\xi_w} - \lambda_\mathcal{L}\left(\int dp|\tilde{\xi}(p)|^2 - 1\right), \tag{3.8}$$

where $\lambda_{\mathcal{L}}$ is the Lagrange multiplier and the constraint condition is the normalization condition for $\tilde{\xi}(p)$. Firstly, we obtain $\tilde{\xi}(p)$ as the solution of the variational problem. Subsequently, we show the expectation value of the initial probe position can be $\langle\hat{x}\rangle_{\xi} = 0$. The final expectation value can be described by Eq. (2.28). Varying the Lagrangian $\mathcal{L}$ with respect to $\lambda_{\mathcal{L}}$, the constraint condition re-emerges as

$$0 = \frac{\partial\mathcal{L}}{\partial\lambda_{\mathcal{L}}} = \int dp|\tilde{\xi}(p)|^2 - 1. \tag{3.9}$$

Varying $\mathcal{L}$ in terms of $\tilde{\xi}^*(p)$, we obtain

$$0 = \frac{\partial\mathcal{L}}{\partial\tilde{\xi}^*} = \frac{i[B^*(p)B'(p)\tilde{\xi}(p) + |B(p)|^2\tilde{\xi}'(p)] - \langle\hat{x}\rangle_{\xi_w}|B(p)|^2\tilde{\xi}(p)}{\int dp|B(p)\tilde{\xi}(p)|^2} - \lambda_{\mathcal{L}}\tilde{\xi}(p). \tag{3.10}$$

This equation means

$$\frac{\tilde{\xi}'(p)}{\tilde{\xi}(p)} = -\frac{B'(p)}{B(p)} - i\left(\langle\hat{x}\rangle_{\xi_w} + \lambda_{\mathcal{L}}|B(p)|^{-2}\int dp|B(p)\tilde{\xi}(p)|^2\right). \tag{3.11}$$

With this equation, the normalization condition (3.9), and the final expectation value $\langle\hat{x}\rangle_{\xi_w}$, we can find $\lambda_{\mathcal{L}} = 0$ as

$$\begin{aligned}\langle\hat{x}\rangle_{\xi_w} &= \frac{i\int dp|B(p)\tilde{\xi}(p)|^2\left(\frac{B'(p)}{B(p)} + \frac{\tilde{\xi}'(p)}{\tilde{\xi}(p)}\right)}{\int dp|B(p)\tilde{\xi}(p)|^2} \\ &= \langle\hat{x}\rangle_{\xi_w} + \lambda_{\mathcal{L}}\int dp|\tilde{\xi}(p)|^2 = \langle\hat{x}\rangle_{\xi_w} + \lambda_{\mathcal{L}}, \ \ \therefore\ \lambda_{\mathcal{L}} = 0.\end{aligned} \tag{3.12}$$

Then, with the indefinite integration over p for Eq. (3.11), we obtain the form of the optimal probe wave function as below:

$$\begin{aligned}\text{Eq. (3.11)} &\Leftrightarrow (\log\tilde{\xi}(p))' = -(\log B(p))' - i\langle\hat{x}\rangle_{\xi_w} \\ &\Rightarrow \int dp(\log\tilde{\xi}(p))' = -\int dp(\log B(p))' - i\int dp\langle\hat{x}\rangle_{\xi_w} \\ &\Rightarrow \log\tilde{\xi}(p) = -\log B(p) - i\langle\hat{x}\rangle_{\xi_w}p + \log C \\ &\Leftrightarrow \tilde{\xi}(p) = CB^{-1}\exp\left[-i\langle\hat{x}\rangle_{\xi_w}p\right],\end{aligned} \tag{3.13}$$

where C is the normalization factor calculated as

$$|C|^2 = \left(\int dp|B(p)|^{-2}\right)^{-1} \tag{3.14}$$

from the normalization condition (3.9) and the integration constant is denoted by $\log C$ for the sake of simplicity. We next evaluate the initial expectation value $\langle\hat{x}\rangle_\xi$ to obtain the shift of the expectation value $\langle\hat{x}\rangle_{\xi_w} - \langle\hat{x}\rangle_\xi$. Substituting Eq. (3.13) into $\langle\hat{x}\rangle_\xi$, we obtain the following equation as

$$\begin{aligned}\langle\hat{x}\rangle_\xi &= \frac{\int dp\langle\xi|p\rangle(i\frac{\partial}{\partial p})\langle p|\xi\rangle}{\int dp\langle\xi|p\rangle\langle p|\xi\rangle} = i\frac{\int dp\tilde{\xi}^*(p)\tilde{\xi}'(p)}{\int dp|\tilde{\xi}(p)|^2} = \langle\hat{x}\rangle_{\xi_w} - i|C|^2\int dp\frac{B^*(p)B'(p)}{|B(p)|^4} \\ &= \langle\hat{x}\rangle_{\xi_w} - g\mathrm{Re}\langle\hat{A}\rangle_w|C|^2\int dp|B(p)|^{-4} - \frac{i}{2}|C|^2\int dp[|B(p)|^{-2}]'. \end{aligned} \tag{3.15}$$

Because the expectation value has to be real-valued and $C \neq 0$, we can find the integration region which meets

$$\int dp[|B(p)|^{-2}]' = 0. \tag{3.16}$$

Focusing on $|B(p)|^{-2}$ in Eq. (2.18) the periodicity of which is $\pi/2g$, we can see that $-\pi/2g \leq p \leq \pi/2g$ would be one of the appropriate region.

With the integration region and Eq. (3.14), we can calculate the normalization factor C as

$$\begin{aligned}|C|^{-2} &= \int_{-\pi/2g}^{\pi/2g} dp|B(p)|^{-2} = \int_{-\pi/2g}^{\pi/2g}\frac{dp}{\cos^2 gp}\frac{1}{1+2\mathrm{Im}\langle\hat{A}\rangle_w\tan gp + |\langle\hat{A}\rangle_w|^2\tan^2 gp} \\ &= \int_{-\infty}^{\infty}\frac{dt}{g}\frac{1}{1+2\mathrm{Im}\langle\hat{A}\rangle_w t + |\langle\hat{A}\rangle_w|^2t^2} \quad \text{substituted } t = \tan gp \\ &= \frac{|\langle\hat{A}\rangle_w|^2}{g(\mathrm{Re}\langle\hat{A}\rangle_w)^2}\int_{-\infty}^{\infty}dt\frac{1}{1+[(\mathrm{Im}\langle\hat{A}\rangle_w + |\langle\hat{A}\rangle_w|^2t)/\mathrm{Re}\langle\hat{A}\rangle_w]^2} \\ &= \frac{|\langle\hat{A}\rangle_w|^2}{g(\mathrm{Re}\langle\hat{A}\rangle_w)^2}\frac{\mathrm{Re}\langle\hat{A}\rangle_w}{|\langle\hat{A}\rangle_w|^2}\left[\arctan\left(\frac{\mathrm{Im}\langle\hat{A}\rangle_w + |\langle\hat{A}\rangle_w|^2t}{\mathrm{Re}\langle\hat{A}\rangle_w}\right)\right]_{-\infty}^{\infty} = \frac{\pi}{g|\mathrm{Re}\langle\hat{A}\rangle_w|}. \end{aligned} \tag{3.17}$$

Furthermore, we can obtain the shift of the expectation value from Eq. (3.15) as

$$\langle\hat{x}\rangle_{\xi_w} - \langle\hat{x}\rangle_\xi = g\mathrm{Re}\langle\hat{A}\rangle_w|C|^2\cdot\frac{|\langle\hat{A}\rangle_w|^2+1}{2(\mathrm{Re}\langle\hat{A}\rangle_w)^2}|C|^{-2} = g\frac{|\langle\hat{A}\rangle_w|^2+1}{2\mathrm{Re}\langle\hat{A}\rangle_w}. \tag{3.18}$$

Here we have calculated the integration:

$$\begin{aligned}
&\int_{-\pi/2g}^{\pi/2g} dp|B(p)|^{-4} = \int_{-\pi/2g}^{\pi/2g} \frac{dp}{\cos^4 gp} \frac{1}{(1+2\mathrm{Im}\langle\hat{A}\rangle_w \tan gp + |\langle\hat{A}\rangle_w|^2 \tan^2 gp)^2} \\
=&\frac{1}{g}\int_{-\infty}^{\infty} dt \frac{1+t^2}{(1+2\mathrm{Im}\langle\hat{A}\rangle_w t + |\langle\hat{A}\rangle_w|^2 t^2)^2} \quad \text{substituted } t = \tan gp \\
=&\frac{|\langle\hat{A}\rangle_w|^2+1}{2(\mathrm{Re}\langle\hat{A}\rangle_w)^2}\int_{-\infty}^{\infty}\frac{dt}{g}\frac{1}{1+2\mathrm{Im}\langle\hat{A}\rangle_w t + |\langle\hat{A}\rangle_w|^2 t^2} \\
&+\frac{1}{2g(\mathrm{Re}\langle\hat{A}\rangle_w)|\langle\hat{A}\rangle_w|^2}\int_{-\infty}^{\infty} dt\left(\frac{\mathrm{Im}\langle\hat{A}\rangle_w(|\langle\hat{A}\rangle_w|^2+1)+[|\langle\hat{A}\rangle_w|^4+|\langle\hat{A}\rangle_w|^2-2(\mathrm{Re}\langle\hat{A}\rangle_w)^2]t}{1+2\mathrm{Im}\langle\hat{A}\rangle_w t+|\langle\hat{A}\rangle_w|^2t^2}\right)' \\
=&\frac{|\langle\hat{A}\rangle_w|^2+1}{2(\mathrm{Re}\langle\hat{A}\rangle_w)^2}|C|^{-2},
\end{aligned} \tag{3.19}$$

where we have used Eq. (3.17) to evaluate the last equal. Here we consider the initial expectation value $\langle\hat{x}\rangle_\xi$. As sated in above, $|B(p)|^{-2}$ has the periodicity such as $|B(p)|^{-2} = |B(p+\pi/g)|^{-2}$ which implies

$$|\tilde{\xi}(p+\pi/g)|^2 = |\tilde{\xi}(p)|^2 \Rightarrow \tilde{\xi}(p+\pi/g) = e^{-i\pi k/g}\tilde{\xi}(p) \Rightarrow \langle\hat{x}\rangle_{\xi_w} = k, \tag{3.20}$$

where k is an arbitrary real constant and indicates the degree of freedom of the phase. Then we can set $\langle\hat{x}\rangle_\xi = 0$ by choosing $k = \langle\hat{x}\rangle_{\xi_w} - \langle\hat{x}\rangle_\xi$.

Consequently, we obtain the optimal probe wave function (3.1) given by Eqs. (3.13), (3.17), and (3.18).

We remark the case that the integration region is m-period, i.e., $-m\pi/2g < p \leq m\pi/2g$ ($m \in \mathbb{N}$) which means

$$|B(m\pi/2g)|^2 = |B(-m\pi/2g)|^2 = \begin{cases} |\langle\hat{A}\rangle_w|^2 & (\text{if } m \text{ is odd.}), \\ 1 & (\text{if } m \text{ is even.}). \end{cases} \tag{3.21}$$

With this integration region, the normalization factor becomes

$$|C_m|^2 = \left(\int_{-m\pi/2g}^{m\pi/2} dp|B(p)|^{-2}\right)^{-1} = \frac{1}{m}\left(\int_{-\pi/2g}^{\pi/2} dp|B(p)|^{-2}\right)^{-1} = \frac{g|\mathrm{Re}\langle\hat{A}\rangle_w|}{m\pi}, \tag{3.22}$$

while the shift of the expectation value becomes

$$\begin{aligned}\langle\hat{x}\rangle_{\xi_w} - \langle\hat{x}\rangle_{\xi} &= g\mathrm{Re}\langle\hat{A}\rangle_w |C_m|^2 \left(\int_{-m\pi/2g}^{m\pi/2} dp|B(p)|^{-4}\right) \\ &= g\mathrm{Re}\langle\hat{A}\rangle_w |C_m|^2 m \frac{|\langle\hat{A}\rangle_w|^2+1}{2(\mathrm{Re}\langle\hat{A}\rangle_w)^2} \frac{\pi}{g|\mathrm{Re}\langle\hat{A}\rangle_w|} = g\frac{|\langle\hat{A}\rangle_w|^2+1}{2\mathrm{Re}\langle\hat{A}\rangle_w}. \end{aligned} \tag{3.23}$$

Then the shift of the probe position expectation value and the amplification factor does not depend on the periodicity of the integration region.

3.3.2 Minimization of the position variance of the final probe distribution

We can derive the optimal probe wave function by deriving the function which provides the minimal variance of final probe wave function, which is given as

$$\begin{aligned} & V[\tilde{\xi}(p), \tilde{\xi}^*(p)] := \langle\hat{x}^2\rangle_{\xi_w} - \langle\hat{x}\rangle_{\xi_w}^2 \\ =& \frac{\int dp[B(p)\tilde{\xi}(p)]^* \left(i\frac{\partial}{\partial p}\right)^2 [B(p)\tilde{\xi}(p)]}{\int dp|B(p)\tilde{\xi}(p)|^2} - \left(\frac{\int dp[B(p)\tilde{\xi}(p)]^* \left(i\frac{\partial}{\partial p}\right) [B(p)\tilde{\xi}(p)]}{\int dp|B(p)\tilde{\xi}(p)|^2}\right)^2 \\ =& -\frac{\int dp[B^*(p)B''(p)|\tilde{\xi}(p)|^2 + 2B^*(p)B(p)'\tilde{\xi}^*(p)\tilde{\xi}'(p) + |B(p)|^2\tilde{\xi}^*(p)\tilde{\xi}''(p)]}{\int dp|B(p)\tilde{\xi}(p)|^2} \\ & - \left(\frac{i\int dp[B^*(p)B'(p)|\tilde{\xi}(p)|^2 + |B(p)|^2\tilde{\xi}^*(p)\tilde{\xi}'(p)]}{\int dp|B(p)\tilde{\xi}(p)|^2}\right)^2. \end{aligned} \tag{3.24}$$

Varying the variance with respect to $\tilde{\xi}^*(p)$, we obtain the following equation as

$$\begin{aligned} 0 =& \frac{\partial V}{\partial\tilde{\xi}^*} \\ =& \frac{-[B^*(p)B''(p)\tilde{\xi}(p) + 2B^*(p)B(p)'\tilde{\xi}'(p) + |B(p)|^2\tilde{\xi}''(p)] - \langle\hat{x}^2\rangle_{\xi_w}|B(p)|^2\tilde{\xi}(p)}{\int dp|B(p)\tilde{\xi}(p)|^2} \\ & - 2\langle\hat{x}\rangle_{\xi_w}\left(\frac{i[B^*(p)B'(p)\tilde{\xi}(p) + |B(p)|^2\tilde{\xi}'(p)] - \langle\hat{x}\rangle_{\xi_m}|B(p)|^2\tilde{\xi}(p)}{\int dp|B(p)\xi(p)|^2}\right) \\ =& \frac{-B(p)^*}{\int dp|B(p)\tilde{\xi}(p)|^2}\left[\left(B(p)\tilde{\xi}(p)\right)'' + 2i\langle\hat{x}\rangle_{\xi_w}\left(B(p)\tilde{\xi}(p)\right)' + (\langle\hat{x}^2\rangle_{\xi_w} - 2\langle\hat{x}\rangle_{\xi_w}^2)B(p)\tilde{\xi}(p)\right]. \end{aligned} \tag{3.25}$$

Therefore it generates the homogeneous linear differential equation of the second order with constant coefficients as

$$0 = X''(p) + 2i\langle\hat{x}\rangle_{\xi_w} X'(p) + (\langle\hat{x}^2\rangle_{\xi_m} - 2\langle\hat{x}\rangle^2_{\xi_m})X(p), \tag{3.26}$$

where $X(p) := B(P)\tilde{\xi}(p)$. We substitute $X(p) = e^{rp}(r \in \mathbb{C})$ into the equation to solve, which produces

$$0 = r^2 + 2i\langle\hat{x}\rangle_{\xi_w} r + \langle\hat{x}^2\rangle_{\xi_m} - 2\langle\hat{x}\rangle^2_{\xi_m}. \tag{3.27}$$

We can easily see the the solutions are $r = -i\langle\hat{x}\rangle_{\xi_w} \pm i\sqrt{V}$. Then we can note the solution of the differential equation with arbitrary constants $C_{1,2}$ as

$$\begin{aligned} X(p) &= C_1 e^{(-i\langle\hat{x}\rangle_{\xi_w} + i\sqrt{V})p} + C_2 e^{(-i\langle\hat{x}\rangle_{\xi_w} - i\sqrt{V})p} \\ &= [(C_1 + C_2)\cos\ \sqrt{V}p + i(C_1 - C_2)\sin\sqrt{V}p]e^{-i\langle\hat{x}\rangle_{\xi_w}p} \\ &= (a\sin\sqrt{V}p + b\cos\sqrt{V}p)e^{-i\langle\hat{x}\rangle_{\xi_w}p}, \end{aligned} \tag{3.28}$$

where $a(:= C_1 + C_2)$ and $b(:= i(C_1 - C_2))$ are the arbitrary complex values, which do not take zero simultaneously. Hereafter we determine the parameters, a, b, V, and the integration region $p_- < p \leq p_+$ (tentatively we use the notation $p_\pm$ to describe the region) for the normalization. We define $Y(p) := a\sin\sqrt{V}p + b\cos\sqrt{V}p$ for convenience. Then we evaluate the final expectation value as

$$\begin{aligned} \langle\hat{x}\rangle_{\xi_w} &= \frac{\int dp X^*(p)\left(i\frac{\partial}{\partial p}\right)X(p)}{\int dp|X(p)|^2} = \frac{i\int dp[e^{-i\langle\hat{x}\rangle_{\xi_w}p}Y(p)]^*[e^{-i\langle\hat{x}\rangle_{\xi_w}p}Y(p)]'}{\int dp|Y(p)|^2} \\ &= \frac{i\int dp\, e^{i\langle\hat{x}\rangle_{\xi_w}p}Y^*(p)[-i\langle\hat{x}\rangle_{\xi_w}e^{-i\langle\hat{x}\rangle_{\xi_w}p}Y(p) + e^{-i\langle\hat{x}\rangle_{\xi_w}p}Y(p)]}{\int dp|Y(p)|^2} \\ &= \langle\hat{x}\rangle_{\xi_w} + \frac{i\int dp Y^*(p)Y'(p)}{\int dp|Y(p)|^2}, \end{aligned} \tag{3.29}$$

therefore we find the parameters satisfy the equation

$$\begin{aligned}
0 &= \int_{p_-}^{p_+} dp Y^*(p) Y'(p) \\
&= \int_{p_-}^{p_+} dp \frac{s}{2}[(|a|^2 - |b|^2)\sin\sqrt{V}p + (ab^* + a^*b)\cos 2\sqrt{V} + (ab^* - a^*b)] \\
&= \int_{p_-}^{p_+} dp \left[\frac{s|a^2+b^2|}{2}\sin(2\sqrt{V}p+\gamma) + i\sqrt{V}\mathrm{Im}ab^*\right] \qquad \left(\gamma := \arctan\frac{2\mathrm{Re}ab^*}{|a|^2-|b|^2}\right) \\
&= \left[-\frac{|a^2+b^2|}{4}\cos(2\sqrt{V}p+\gamma) + i\sqrt{V}\mathrm{Im}ab^* p\right]_{p_-}^{p_+} \\
&= -\frac{|a^2+b^2|}{4}\left[\cos(2\sqrt{V}p_+ + \gamma) - \cos(2\sqrt{V}p_- + \gamma)\right] + i\sqrt{V}\mathrm{Im}ab^*(p_+ - p_-). \qquad (3.30)
\end{aligned}$$

We can choose the parameters to satisfy this equation in three ways as follows;

Case 1. The variance is zero ($V = 0$). In this case b becomes the normalization factor. Hence, the wave function becomes the optimal probe as derived in the previous section.

Case 2. $V \neq 0$, $a = ib$, and $\mathrm{Im}ab^* = 0$. They give $a = b = 0$ which is not appropriate for the current interest.

Case 3. $V \neq 0$, $\cos(2\sqrt{V}p_+ + \gamma) = \cos(2\sqrt{V}p_- + \gamma)$, and $\mathrm{Im}ab^* = 0$. The second condition provides

$$\begin{aligned}
&-2\sin(\sqrt{V}p_+ + \sqrt{V}p_- + \gamma)\sin(\sqrt{V}p_+ - \sqrt{V}p_-) = 0, \\
&\therefore\ \sqrt{V}p_+ + \sqrt{V}p_- + \gamma = n\pi \text{ or } \sqrt{V}p_+ - \sqrt{V}p_- = n\pi.
\end{aligned}$$

Then we have found that the form of the optimal probe wave function can give the minimal variance, i.e., the zero variance (Case 1). Subsequent calculations for the normalization factor, the integration region, and the shift of the expectation value, are the same as the previous section. Here we omit these calculations.

3.4 Summary of this chapter

In this chapter, we have derived the optimal probe wave function by the Lagrange multiplier method. Ref. [35] shows that the amplification limit in the weak interaction case

and any number of the distinct eigenvalues of the measured system. Here we have calculated the optimal probe and the amplified shift without any approximation, where the measured system is the two-state system. The optimal wave function is described in the momentum space as Eq. (3.1), the support of which is $-\pi/2g \leq p \leq \pi/2g$. This support can be extended the region $-nm/2g \leq p \leq nm/2g\,(m \in \mathbb{N})$, which leads to the same conclusion. The weak measurement with the wave function gives the shift of the expectation value of the probe position as Eq. (3.2), which has no upper bound as $|\langle \hat{A} \rangle_w|$ becomes large. Furthermore when we choose such a weak value that gives the final probe position expectation value $\langle \hat{x} \rangle_{\xi_w} = 2gk\,(k \in \mathbb{N})$, the final probe position distribution becomes the Kronecker delta $\xi_w(x = 2gj) = \delta_{jk}$ for $j \in \mathbb{N}$, i.e., the variance of which can be zero. Namely, we can extract the value of the coupling constant from the obtained data without the statistical fluctuation in principle. Here we remark that when the coupling constant is sufficiently small, the support of the function covers the whole momentum space. Then, the final probe wave function in the position space behaves like a delta function. To derive the wave function, we assume that the observable $\hat{A}$ of the measured system satisfies $\hat{A}^2 = 1$ and the weak value $\mathrm{Re}\langle \hat{A} \rangle_w \neq 0$ which is for the normalization of the statically.

Note that there is another wave function which can give an arbitrary amplification factor, for example, the wave function proposed in Ref. [62] [2] . Anyhow, the essence of the arbitrary amplification is that the denominator of the wave function (3.1) denoted by $B(x)$ cancels out the higher-order terms given by the unitary operator which brings the upper bound of amplification. Among such wave functions, the optimal probe wave function is the only solution that can decrease the variance of the final position distribution to zero.

We should pay attention to engineering the wave function that we need to use the coupling constant g and the weak value $\langle \hat{A} \rangle_w$. We can grasp the weak value from the chosen pre- and postselected states in a given experimental setup. On the other hand the value of the coupling constant g is unknown obtainable only by the experiment. Then, initially, we need to choose the value by a reasonable guess for construction of a tentative probe wave function. From the discrepancy of the coupling constant from the theoretical prediction and the actual experimental data, the value of the coupling constant tends to a more and more accurate value by iteration.

[2] The proposed wave function is $\tilde{\xi}_i(p) = e^{-\alpha G(p)}/B(p)$, the support of which is $-n\pi \leq gp \leq n$ for $n \in \mathbb{N}$. Here α is an arbitrary real number and $G(p)$ is a primitive of $|B(p)|^{-2}$. This wave function gives the initial and final position expectation value as $\langle \hat{x} \rangle_i = g(\alpha - \mathrm{Re}\langle \hat{A} \rangle_w)(1 + |\langle \hat{A} \rangle_w|^2)/2(\mathrm{Re}\langle \hat{A} \rangle_w)^2$ and $\langle \hat{x} \rangle_f = -g\alpha/|\mathrm{Re}\langle \hat{A} \rangle_w|$. Then the shit of the expectation value is $\Delta\langle \hat{x} \rangle/g = (1 + |\langle \hat{A} \rangle_w|^2)/2\mathrm{Re}\langle \hat{A} \rangle_w + \alpha[(1 - |\mathrm{Re}\langle \hat{A} \rangle_w|)^2 + (\mathrm{Im}\langle \hat{A} \rangle_w)^2]/2(\mathrm{Re}\langle \hat{A} \rangle_w)^2$.

Chapter 4

Review of Statistical Hypothesis Testing

4.1 Preface

In this chapter, we recaptulate the concept of the standard hypothesis testing established by Jerzy Neyman and Egon Sharpe Pearson [63]. The hypothesis testing is one of the statistical inference methods, which is widely used not only in physical experiments but also in other research fields [53–55]. The hypothesis testing method provides the mathematical decision for the testing problem to judge which hypothesis is more plausible in two contradictory hypotheses. For the good decision-making, we have to produce the appropriate testing method corresponding to a problem. Here we consider the uniformly most powerful (UMP) test, and the uniformly most powerful unbiased (UMPU) test. To write this chapter, the author have refered to Ref. [45].

4.2 Definitions of technical terms and notations

Here we introduce some technical terminologies and notations. In the hypothesis testing, we attempt to determine which hypothesis is more proper in two contradictory hypotheses, i.e., a null hypothesis and an alternative hypothesis.

Table 4.1: The relation between our decision and the actual hypothesis.

		actual hypothesis	
		null	alternative
our decision	null	correct	the type-2 error
	alternative	the type-1 error	correct

Here we consider what the value of the unknown real parameter θ is[1]. We define the two parameter sets, Θ_0 and Θ_1 such that $\Theta_0 \cap \Theta_1 = \{\phi\}$. We call the null hypothesis H_0 if the parameter θ is included in Θ_0, i.e., $H_0 : \theta \in \Theta_0$. Similarly the alternative one is $H_1 : \theta \in \Theta_1$. Hereafter, we simply describe the testing problem as

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta_1. \tag{4.1}$$

If the set Θ is a single point set such as $\Theta = \{\theta_0\}$, the hypothesis is called the simple hypothesis. Meanwhile, the hypothesis that is not simple is classified as the composite hypothesis. In this chapter, we consider the three testing problems below.

1. the both hypotheses are simple, i.e., $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$,
2. the one-sided test, i.e., $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$,
3. the two-sided test, i.e., $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$.

Sometimes the testing problems such as $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta_0$ and $H_0 : \theta = \theta_0$ vs. $H_1 : \theta < \theta_0$ are classified as the one-sided test. We state "we accept the null (alternative) hypothesis," when we determine that the null (alternative) hypothesis is correct. On the other hand, we state "we reject the null (alternative) hypothesis," when we determine that the null (alternative) hypothesis is incorrect.

Next we consider a cost or risk that we have to pay in the process which we want to evaluate. In the standard hypothesis testing theory, we regard the error, i.e., the misjudging which hypothesis is true as the cost. We have two types of errors as

type-1 error: we wrongly reject the null hypothesis, whereas the null one is correct,

[1]In this thesis, we assume that θ is an one dimensional parameter, i.e., $\theta \in \mathbb{R}$.

type-2 error: we wrongly accept the null hypothesis, whereas the alternative one is correct.

Obviously, a good testing is the one that the probabilities of these two errors are small, which are calculated from a distribution function. Our objective in the hypothesis testing is to find such a decision-making procedure with small errors. However, generally, it is very difficult to make the two errors simultaneously small. Hence, we are content to take the standard strategy that we make the probability of the type-2 error small, while we control the probability of the type-1 error to be smaller than a certain value called a significance level $\alpha(0 \leq \alpha \leq 1)$ [54]. This strategy is called as the Neyman-Pearson Criterion.

To calculate these probabilities, we choose a test function[2] $d(x)$ which depends on the obtained data x by an experiment. The test function $d(x)$ takes the binary values of 0 or 1. The 0 indicates that we accept the null hypothesis, and the 1 represents that we accept the alternative one. We can describe the test function as

$$d(x) = \begin{cases} 0 & \text{for } T(x) < c, \\ r & \text{for } T(x) = c, \\ 1 & \text{for } T(x) > c, \end{cases} \tag{4.2}$$

where $T(x)$ is a test statistics which indicates the decision criteria with the obtained data x, and c is a critical point which gives the significance level. Theoretically, we can take an arbitrarily real value as c. The region of x satisfying $T(x) > c$ is a rejection region (or a critical region), which indicates the region of rejecting the null hypothesis. Similarly the region of x such as $T(x) < c$ is an acceptance region. When $T(x) = c$, we randomly accept the null hypothesis with the probability $r(0 \leq r \leq 1)$. For the distribution function of continuous variable, the probability for the continuous random variable x coincides just with the case $T(x) = c$ is zero. An appropriate test statistics $T(x)$ depends on a testing problem.

With the decision function $d(x)$, we can define a power of a statistical test as

$$\beta_d(\theta) := \int d(x) f(x|\theta) dx, \tag{4.3}$$

where $f(x|\theta)$ is the probability distribution x with the parameter θ. The test function for all $\theta \in \Theta_1$ represents the probability that the accepted alternative hypothesis is correct,

[2]This corresponds to the decision function in the statistical interference. In the hypothesis testing function, we usually use the term "test function" instead of "decision function."

while the test function for all $\theta \in \Theta_0$ is the probability of the type-1 error. Here the probability is calculated by the integration of the distribution function $f(x|\theta)$ for $\theta \in \Theta_0$ over the rejection region. We can evaluate the probability of the type-1 error $\Pr[\mathcal{E}_1]$ as

$$\Pr[\mathcal{E}_1] \equiv \beta_d(\theta) \text{ for } \theta \in \Theta_0. \tag{4.4}$$

Also the probability of the type-2 error $\Pr[\mathcal{E}_2]$ is calculated by the integration of the distribution function $f(x|\theta)$ for $\theta \in \Theta_1$ over the acceptance region, which is given as

$$\Pr[\mathcal{E}_2] \equiv 1 - \beta_d(\theta) \text{ for } \theta \in \Theta_1. \tag{4.5}$$

Therefore, we can obtain the probabilities of the type-1 and -2 errors. As we stated in the above, the smaller the probabilities of errors are, the better the decision-making becomes. It is important to appropriately choose the test function because it gives the probabilities of errors. In the next section, we propose a proper test function for one testing problem.

4.3 Uniformly most powerful (UMP) test

As we noted in the previous section, we customarily take the strategy such that we reduce the probability of the type-2 error, while we keep the probability of the type-1 error smaller than a certain significance point. In this strategy, we can say that the best test method is a uniformly most powerful (UMP) test.

Definition 4.3.1. (Uniformly most powerful test)
We assume that the testing problem is $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$ and $d(x)$ is an arbitrary test function ($0 \leq d(x) \leq 1$) which gives the test at a significance level α for the testing problem, namely $\beta_d(\theta) \leq \alpha$ is satisfied for all $\theta \in \Theta_0$. If the test function $d^*(x)$, which gives the test at the significance level α, satisfies the inequality

$$\beta_{d^*}(\theta) \geq \beta_d(\theta), \ {}^\forall\theta \in \Theta_1, \tag{4.6}$$

the test is UMP.

Namely, the test gives the maximum power for all alternative hypotheses (${}^\forall\theta \in \Theta_1$). When the alternative hypothesis is simple, a test satisfying the inequality (4.6) is simply

called most powerful (MP) test. We remark that we cannot always find a UMP test or a MP test in every testing problem. Precisely, there is a testing problem such that a UMP test or an MP test is not present.

Here we exemplify a MP test for the testing problem such that the both hypotheses are simple, which is addressed by Neyman-Pearson's lemma [63]. This lemma claims that the test named likelihood-ratio test is MP only when the both hypotheses are simple.

Lemma 4.3.1. (Neyman-Pearson's lemma)
We assume that the testing problem is $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$, and that $f(x|\theta_0)$ and $f(x|\theta_1)$ are the probability distributions in the null hypothesis and the alternative one, respectively. We consider the following test function

$$d^*(x) = \begin{cases} 0 & \text{for } f(x|\theta_1)/f(x|\theta_0) < c, \\ r & \text{for } f(x|\theta_1)/f(x|\theta_0) = c, \\ 1 & \text{for } f(x|\theta_1)/f(x|\theta_0) > c, \end{cases} \tag{4.7}$$

in which the probability of the type-1 error is α, i.e., $\beta_{d^*}(\theta_0) = \alpha$. This test function gives the MP test at the significance level α for the hypothesis testing problem.

The $f(x|\theta_1)/f(x|\theta_0)$ is called the likelihood-ratio. Here we give the proof of the lemma.

Proof. We set that $d(x)$ is an arbitrary test function at a significance level α, namely $\beta_d(\theta_0) \leq \alpha$ and $0 \leq d(x) \leq 1$. First of all, we show the inequality

$$\int [d^*(x) - d(x)][f(x|\theta_1) - cf(x|\theta_0)]dx \geq 0. \tag{4.8}$$

From Eq. (4.7) and $0 \leq d(x) \leq 1$, we can see

$$d^*(x) - d(x) \geq 0 \text{ for } f(x|\theta_1) - cf(x|\theta_0) > 0, \tag{4.9}$$

$$d^*(x) - d(x) \leq 0 \text{ for } f(x|\theta_1) - cf(x|\theta_0) < 0. \tag{4.10}$$

Additionally, the left-hand side of the inequality (4.8) is zero when $f(x|\theta_1) - cf(x|\theta_0) = 0$. Thus the inequality (4.8) holds, because the integrand is non-negative for all x. Therefore

we can show

$$\begin{aligned}\beta_{d^*}(\theta_1) - \beta_d(\theta_1) &= \int d^*(x) f(x|\theta_1) dx - \int d(x) f(x|\theta_1) dx \\ &= \int [d^*(x) - d(x)] f(x|\theta_1) dx \\ &\geq c \int [d^*(x) - d(x)] f(x|\theta_0) dx \\ &= c \left[\int d^*(x) f(x|\theta_0) dx - \int d(x) f(x|\theta_0) dx \right] \\ &= c \left[\alpha - \beta_d(\theta_0) \right] \geq 0 \end{aligned} \tag{4.11}$$

Thus we have shown that the likelihood-ratio test is the MP test. □

We have to pay attention to the fact that this lemma holds only when the both hypotheses are simple. Subsequently, we can find a UMP test also in the one-sided test. Prior to showing it, we define the monotone likelihood-ratio.

Definition 4.3.2. (Monotone likelihood ratio)
Let us consider the likelihood-ratio $f(x|\theta_2)/f(x|\theta_1)$ for arbitrarily fixed θ_1 and θ_2 ($\theta_1 < \theta_2$). We assume that the likelihood-ratio can be described as

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = g(T(x), \theta_1, \theta_2). \tag{4.12}$$

When $g(T(x), \theta_1, \theta_2)$ is a monotonically increasing function of $T(x)$ for $\theta_1 < \theta_2$, we state that the distribution function $f(x|\theta)$ has a monotone likelihood ratio in $T(x)$.

We see the theorem to find a UMP test in the one-sided test.

Theorem 4.3.1. (Uniformly most powerful test in the one-sided test)
Here we consider the testing problem $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$. When the distribution function $f(x|\theta)$ has a monotone likelihood-ratio in the statistics $T(x)$, the test function

$$d^*(x) = \begin{cases} 0 & \text{for } T(x) < c, \\ r & \text{for } T(x) = c, \\ 1 & \text{for } T(x) > c, \end{cases} \tag{4.13}$$

becomes a UMP test based on the critical point c and the random probability r for an arbitrary significance level α ($0 \leq \alpha \leq 1$).

This theorem is even valid for the testing problem $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta_0$. The proof is the following.

Proof. Firstly, we consider the particular testing problem $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ with arbitrary fixed θ_0 and θ_1 ($\theta_1 > \theta_0$). As shown in Neyman-Pearson's lemma, the likelihood-ratio test becomes MP test, which can be described as the form of Eq. (4.13). The critical point c and the random probability r are independent of the parameter θ_1. The test function (4.13) gives the UMP test for the testing problem $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta_0$.

Next we expand the null hypothesis $H_0 : \theta = \theta_0$ into $\theta \leq \theta_0$. In preparation, we show the following inequality

$$\int g(T(x))f(x)dx \int h(T(x))f(x)dx \leq \int g(T(x))h(T(x))f(x)dx, \tag{4.14}$$

where $g(T(x))$ and $h(T(x))$ are the monotonically increasing functions of $T(x)$. Setting $m = \int g(T(x))f(x)dx$, we can obtain the equation

$$0 = \int [g(T(x)) - m]f(x)dx \tag{4.15}$$

due to the probability distribution $f(x)$. Since $g(T(x))-m$ is the monotonically increasing function of $T(x)$, there is a finite value t such as $g(T(x)) - m \leq 0$ for $T(x) \leq t$ and $g(T(x)) - m > 0$ for $T(x) > t$ to satisfy Eq. (4.15). $h(T(x))$ is also the monotonically increasing function of $T(x)$, we can obtain the inequalities, $h(T(x))-h(t) \leq 0$ for $T(x) \leq t$ and $h(T(x)) - h(t) > 0$ for $T(x) > t$. Therefore, the inequality

$$[g(T(x)) - m][h(T(x)) - h(t)] \leq 0 \tag{4.16}$$

holds for all x, which indicates

$$\begin{aligned}
0 &\leq \int [g(T(x)) - m][h(T(x)) - h(t)]f(x)dx \\
&= \int [g(T(x)) - m]h(T(x))f(x)dx - h(t) \int [g(T(x)) - m]f(x)dx \\
&= \int g(T(x))h(T(x))f(x)dx - m \int h(T(x))f(x)dx \quad (\because \text{Eq.}(4.15)) \\
&= \int g(T(x))h(T(x))f(x)dx - \int g(T(x))f(x)dx \int h(T(x))f(x)dx.
\end{aligned} \tag{4.17}$$

Hence we have shown Eq. (4.14). Here we assume $\theta \leq \theta_0$ to evaluate the the probability of the type-1 error and . In Eq. (4.14), substituting $f(x) = f(x|\theta)$, $g(T(x)) = f(x|\theta_0)/f(x|\theta)$, and $h(T(x)) = d^*(x)$, we obtain

$$\begin{aligned}
&\int \frac{f(x|\theta_0)}{f(x|\theta)} f(x|\theta)dx \int d^*(x)f(x|\theta)dx \leq \int \frac{f(x|\theta_0)}{f(x|\theta)} d^*(x)f(x|\theta)dx \\
&\Leftrightarrow \int d^*(x)f(x|\theta)dx \leq \int f(x|\theta_0)d^*(x)dx \\
&\Leftrightarrow \beta_{d^*}(\theta) \leq \beta_{d^*}(\theta_0) = \alpha
\end{aligned} \tag{4.18}$$

For $\theta \leq \theta_0$, the probability of the type-1 error is smaller than the significance level α. Therefore the test function $d^*(x)$ gives a UMP test for the one-sided test $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$ and $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta_0$. □

In this section, we see a likelihood-ratio test and a UMP test for the specific testing problems, the case that the both hypotheses are simple and the one-sided test, respectively. In the next section, we consider a better test function in a two-sided test.

4.4 Uniformly most powerful unbiased (UMPU) test

We have introduced the UMP test for a one-sided test in the previous section. However, a UMP test does not exist in a two-sided test. We note this issue in Sec. 4.5. Here we find the better test function in a two-sided test. Firstly, we introduce the unbiased test to find one.

Definition 4.4.1. (Unbiased test)
In the testing problem $H_0 : \theta \in \Theta$ vs. $H_1 : \theta \in \Theta$, if the test function $d(x)$ at the significance level α satisfies

$$\beta_d(\theta) \geq \alpha, \ ^\forall\theta \in \Theta_1. \tag{4.19}$$

we state that the test is unbiased.

Next we define the class of a better testing in the unbiased test. Similar to the UMP test, the unbiased test which gives the maximum power for all alternative hypotheses is preferable.

Definition 4.4.2. (Uniformly most powerful unbiased test)
We assume that the testing problem is $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$ and $d(x)$ is an arbitrary unbiased test function at a significance level α, namely $\beta_d(\theta) \leq \alpha$ for all $\theta \in \Theta_0$ and $0 \leq d(x) \leq 1$. If the unbiased test function $d^*(x)$, which gives a test at the significance level α, satisfies the inequality

$$\beta_{d^*}(\theta) \geq \beta_d(\theta), \ {}^{\forall}\theta \in \Theta_1, \tag{4.20}$$

we state that the test function $d^*(x)$ gives uniformly most powerful unbiased (UMPU) test.

Hereafter we consider the unbiased test at the significance level α, i.e., $\beta_d(\theta_0) \leq \alpha$ for the two-sided test $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$. We assume that the power function $\beta_d(x)$ is differentiable with respect to θ and satisfies

$$\frac{\partial}{\partial\theta}\beta_d(\theta) = \int d(x)\frac{\partial}{\partial\theta}f(x|\theta)dx. \tag{4.21}$$

This assumption implies that $\beta_d(\theta)$ is the continuous function of θ. By the unbiasedness of the test, $\beta_d(\theta) \geq \alpha$ for all $\theta \in \Theta_1$, and the continuity of the power function $\beta_d(\theta)$ at $\theta = \theta_0$, we can deduce that the probability of the type-1 error is fixed α, i.e., $\beta_d(\theta_0) = \alpha$. Furthermore, it is clear that the power function $\beta_d(\theta)$ takes minimum value at $\theta = \theta_0$, which implies $\beta'_d(\theta_0) = 0$. The prime indicates the differential with respect to θ. For the unbiased test with these properties, we have the following lemma to obtain a test function giving a UMPU test in the two-sided testing problem.

Lemma 4.4.1. (Uniformly most powerful unbiased test)
We consider the two-sided testing problem $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$. We assume that the test function $d^*(x)$ satisfies

$$\frac{\partial}{\partial\theta}\beta_{d^*}(\theta) = \int d^*(x)\frac{\partial}{\partial\theta}f(x|\theta)dx, \tag{4.22}$$

$$\beta_{d^*}(\theta_0) = \alpha, \tag{4.23}$$

$$\beta'_{d^*}(\theta_0) = 0. \tag{4.24}$$

For an arbitrary fixed $\theta_1(\neq \theta_0)$, when the test function can be described as

$$d^*(x) = \begin{cases} 0 & \text{for } \mathcal{F}(x) < 0, \\ r & \text{for } \mathcal{F}(x) = 0, \\ 1 & \text{for } \mathcal{F}(x) > 0, \end{cases} \tag{4.25}$$

where

$$\mathcal{F}(x) = f(x|\theta_1) - c_1 f(x|\theta_0) - c_2 \frac{\partial}{\partial\theta} f(x|\theta_0), \tag{4.26}$$

for specific c_1 and c_2, the test function $d^*(x)$ gives a unbiased test, which maximizes the power at $\theta = \theta_1$. Furthermore, for the arbitrary $\theta_1(\neq \theta_0)$, if we can appropriately choose $c_1 = c_1(\theta_1)$ and $c_2 = c_2(\theta_1)$, the test function $d^*(x)$ gives a UMPU test.

Proof. The proof is similar to the one of the Nyeman-Pearson Lemma. We set an arbitrarily unbiased test function $d(x)$ $(0 \leq d(x) \leq 1)$ at the significant α, which meets

$$\beta_d(\theta_0) \leq \alpha \quad \text{and} \quad \beta_d(\theta_1) \geq \alpha, \ {}^\forall\theta_1 \neq \theta_0. \tag{4.27}$$

Also we assume that the power function $\beta_d(\theta)$ of the the test function $d(x)$ satisfies Eq. (4.21), which implies that

$$\beta_d(\theta_0) = \alpha \quad \text{and} \quad \frac{\partial}{\partial\theta}\beta_d(\theta_0) = \int d(x) \frac{\partial}{\partial\theta} f(x|\theta_0) dx = 0. \tag{4.28}$$

First we show the inequality

$$\int [d^*(x) - d(x)]\mathcal{F}(x) dx \geq 0. \tag{4.29}$$

From Eqs. (4.25), (4.26) and $0 \leq d(x) \leq 1$, we can see

$$d^*(x) - d(x) \geq 0 \text{ for } \mathcal{F}(x) > 0, \tag{4.30}$$

$$d^*(x) - d(x) \leq 0 \text{ for } \mathcal{F}(x) < 0. \tag{4.31}$$

Additionally, when $\mathcal{F}(x) = 0$, the integrand of Eq. (4.29) is zero regardless of the sing of $d^*(x) - d(x)$. Therefore the inequality (4.29) holds for all x. Next, we see the left-hand

side of the inequality (4.29), which is evaluated as

$$\begin{aligned}
&\int [d^*(x) - d(x)]\mathcal{F}(x) \\
&= \int d^*(x) f(x|\theta_1) dx - \int d(x) f(x|\theta_1) dx \\
&\quad - c_1 \left(\int d^*(x) f(x|\theta_0) dx - \int d(x) f(x|\theta_0) dx \right) \\
&\quad - c_2 \left(\int d^*(x) \frac{\partial}{\partial\theta} f(x|\theta_0) dx - \int d(x) \frac{\partial}{\partial\theta} f(x|\theta_0) dx \right) \\
&= \beta_{d^*}(\theta_1) - \beta_d(\theta_1) - c_1(\beta_{d^*}(\theta_0) - \beta_d(\theta_0)) - c_2 \left(\frac{\partial}{\partial\theta}\beta_{d^*}(\theta_0) - \frac{\partial}{\partial\theta}\beta_d(\theta_0) \right) \\
&= \beta_{d^*}(\theta_1) - \beta_d(\theta_1) - c_1(\alpha - \alpha) - c_2(0 - 0) \\
&= \beta_{d^*}(\theta_1) - \beta_d(\theta_1). \qquad (4.32)
\end{aligned}$$

To show this equation, we have used Eqs. (4.22), (4.23), (4.24), and (4.28). From Eqs. (4.27), (4.29), and (4.32), we can obtain the inequality

$$\beta_{d^*}(\theta) \geq \beta_d(\theta) \geq \alpha, \ {}^\forall\theta \in \Theta_1, \qquad (4.33)$$

because θ_1 is the arbitrary value except for θ_0.

Hence $d^*(x)$ gives the UMPU test for the two-sided test $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$. □

4.5 Examples of the UMP test and UMPU test

In this section, we see the UMP test and UMPU test by exemplifying the testing problem about a mean value θ of Gaussian distribution given as $f(x|\theta) = e^{-(x-\theta)^2/(2\sigma^2)}/\sqrt{2\pi\sigma^2}$. Here we set the two testing problems: the one-sided test $H_0 : \theta = 0$ vs. $H_1 : \theta > 0$ (and also $H_0 : \theta \leq 0$ vs. $H_1 : \theta > 0$) to consider the UMP test and the two-sided test $H_0 : \theta = 0$ vs. $H_1 : \theta \neq 0$ to consider the UMPU test.

Here we see the one-sided test. To obtain the UMP test, we calculate the likelihood-ratio as

$$\frac{f(x|\theta)}{f(x|0)} = \frac{e^{-\frac{(x-\theta)^2}{2\sigma^2}}}{e^{-\frac{x^2}{2\sigma^2}}} = e^{-\frac{\theta^2}{2\sigma^2}} \exp\left[\frac{x\theta}{\sigma^2}\right]. \qquad (4.34)$$

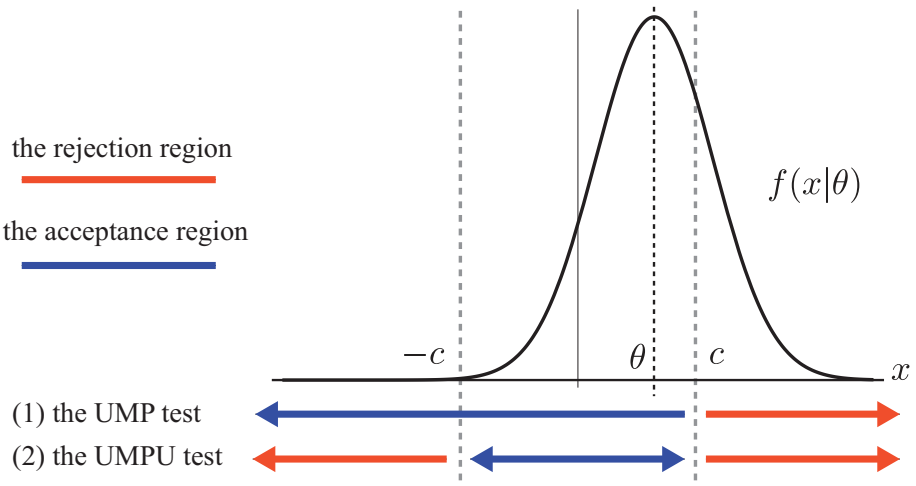


Figure 4.1: The illustration of the rejection region and the acceptance region for (1) the UMP test and (2) the UMPU test.

Setting the test statistics $T(x) = x$, we can find that the likelihood-ratio is the monotonically increasing function of $T(x)$. Hence we propose the test function for the one-sided test [Fig. 4.1(1)] as

$$d_1(x) = \begin{cases} 0 & \text{for } x < c, \\ r & \text{for } x = c, \\ 1 & \text{for } x > c, \end{cases} \tag{4.35}$$

which gives the power calculated as

$$\beta_{d_1}(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_c^\infty e^{-\frac{(x-\theta)^2}{2\sigma^2}} = \frac{1}{2} - \frac{1}{\sqrt{\pi}} \int_0^{\frac{c-\theta}{\sqrt{2\sigma^2}}} e^{-x^2} dx = \frac{1}{2}\left(1 - \text{erf}\left[\frac{c-\theta}{\sqrt{2\sigma^2}}\right]\right), \tag{4.36}$$

where

$$\text{erf}\,[x] := \frac{2}{\sqrt{\pi}} \int_0^x dt e^{-t^2} \tag{4.37}$$

is the error function[3]. Now we consider the two-sided test, in which the appropriately test function is given by the Lemma 4.4.1. To find the UMPU test, we firstly look at the candidate of the UMPU test and check that the candidate meets the Lemma. Here we

[3]The error function satisfies $\text{erf}\,[-x] = -\text{erf}\,[x]$ and $\frac{d}{dx}\text{erf}\,[x] = \frac{2}{\sqrt{\pi}} e^{-x^2}$.

propose the test function [Fig. 4.1(2)] as

$$d_2(x) = \begin{cases} 0 & \text{for } |x| < c, \\ r & \text{for } |x| = c, \\ 1 & \text{for } |x| > c, \end{cases} \tag{4.38}$$

which gives the power

$$\begin{aligned} \beta_{d_2}(\theta) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{-c} e^{-\frac{(x-\theta)^2}{2\sigma^2}} dx + \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-c}^{-\infty} e^{-\frac{(x-\theta)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{\pi}} \left(\int_{\frac{c+\theta}{\sqrt{2\sigma^2}}}^{\infty} e^{-x^2} dx + \int_{\frac{c-\theta}{\sqrt{2\sigma^2}}}^{\infty} e^{-x^2} dx \right) \\ &= 1 - \frac{1}{2} \left(\text{erf} \left[\frac{c-\theta}{\sqrt{2\sigma^2}} \right] + \text{erf} \left[\frac{c+\theta}{\sqrt{2\sigma^2}} \right] \right). \end{aligned} \tag{4.39}$$

We can confirm that Eqs. (4.22), (4.23) and (4.24) are satisfied as

$$\frac{\partial}{\partial\theta} \beta_{d_2}(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} (e^{-\frac{(c-\theta)^2}{\sqrt{2\sigma^2}}} - e^{-\frac{(c+\theta)^2}{\sqrt{2\sigma^2}}}), \tag{4.40}$$

$$\begin{aligned} \int d_2(x) \frac{\partial}{\partial\theta} f(x|\theta) dx &= \int d_2(x) \frac{(x-\theta)}{\sigma^2} f(x|\theta) dx \\ &= \frac{1}{\sigma^2} \frac{1}{\sqrt{2\pi\sigma^2}} \left(\int_{-\infty}^{-c} (x-\theta) e^{-\frac{(x-\theta)^2}{2\sigma^2}} dx + \int_{c}^{\infty} (x-\theta) e^{-\frac{(x-\theta)^2}{2\sigma^2}} dx \right) \\ &= \frac{2}{\sqrt{2\pi\sigma^2}} \left(- \int_{\frac{c+\theta}{\sqrt{2\sigma^2}}}^{\infty} x e^{-x^2} dx + \int_{\frac{c-\theta}{\sqrt{2\sigma^2}}}^{\infty} x e^{-x^2} dx \right) \\ &= \frac{2}{\sqrt{2\pi\sigma^2}} \left(- \left[-\frac{1}{2} e^{-x^2} \right]_{\frac{c+\theta}{\sqrt{2\sigma^2}}}^{\infty} + \left[-\frac{1}{2} e^{-x^2} \right]_{\frac{c-\theta}{\sqrt{2\sigma^2}}}^{\infty} \right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} (e^{-\frac{(c-\theta)^2}{\sqrt{2\sigma^2}}} - e^{-\frac{(c+\theta)^2}{\sqrt{2\sigma^2}}}), \end{aligned} \tag{4.41}$$

$$\beta_{d_2}(0) = 1 - \text{erf} \left[\frac{c}{\sqrt{2\sigma^2}} \right], \tag{4.42}$$

and $\beta'_{d_2}(0) = 0$ is obvious by Eq. (4.40). In Eq.(4.42), by choosing a proper c, we have $\beta_{d_2}(0) = \alpha$.

Next we choose the adequate parameters c_1 and c_2. We evaluate Eq. (4.26) as

$$\frac{\mathcal{F}(x)}{f(x|0)} = e^{\frac{x\theta-\theta^2}{2\sigma^2}} - c_1 - c_2 \frac{x}{\sigma^2}. \tag{4.43}$$

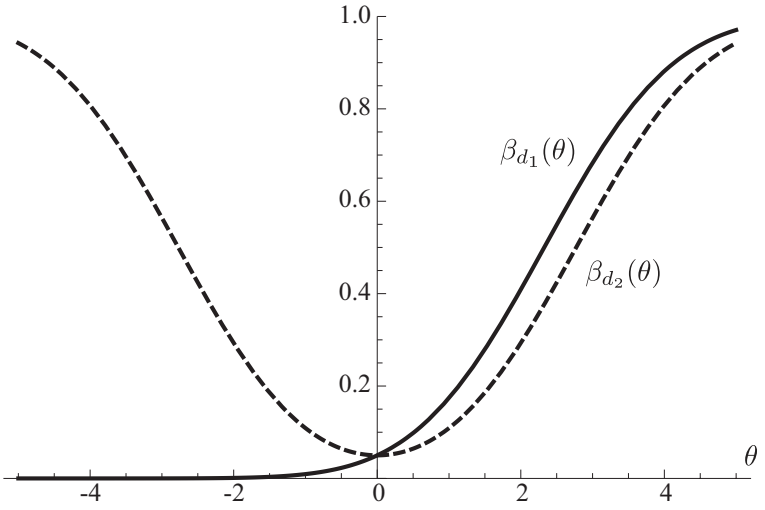


Figure 4.2: Plots of the powers of $\beta_{d_1}(\theta)$ given by the UMP test and $\beta_{d_2}(\theta)$ by the UMPU test. The significance level is fixed as $\alpha = 0.05$ and the critical points c are tuned for which, respectively. The variance of the Gaussian probe is fixed as $\sigma = 2$.

We can find c_1 and c_2 which gives $x = \pm c$ as the solutions of $\mathcal{F}(x)/f(x|0) = 0$ below,

$$c_1 = \frac{e^{c\theta-\theta^2} + e^{-c\theta-\theta^2}}{2} = e^{-\theta^2}\cosh c\sigma, \tag{4.44}$$

$$c_2 = \frac{e^{c\theta-\theta^2} - e^{-c\theta-\theta^2}}{2c/\sigma^2} = \frac{\sigma e^{-\theta^2}}{c}\sinh c\sigma. \tag{4.45}$$

Due to the convexity of the function $e^{x\theta-\theta^2}$, the test function gives UMPU test because

$$d(x) = \begin{cases} 0 & \text{for } \mathcal{F}(x) < 0 \Leftrightarrow |x| < c, \\ r & \text{for } \mathcal{F}(x) = 0 \Leftrightarrow |x| = c, \\ 1 & \text{for } \mathcal{F}(x) > 0 \Leftrightarrow |x| > c. \end{cases} \tag{4.46}$$

Hence we have found the UMP and the UMPU tests for the one-sided test and the two-sided test, respectively. Here we see the difference of the powers between $\beta_{d_1}(x)$ and $\beta_{d_2}(x)$.

Figure 4.2 displays the plots of the powers $\beta_{d_1}(\theta)$ and $\beta_{d_2}(\theta)$, which show that the UMP test for the one-sided test gives the power such as $\beta_{d_1}(\theta \leq 0) \leq \alpha$, while the UMPU test for the two-sided test gives the power $\beta_{d_2}(\theta = 0) \leq \alpha$ and $\beta_{d_2}(\theta \neq 0) \geq \alpha$ as stated in the previous sections. In each test, the powers are lower than the significance level α at the respective null hypotheses. Furthermore, we can see $\beta_{d_1}(\theta) > \beta_{d2}(\theta)$ in the region of

$\theta > 0$[4].

We remark the reason why a UMP test does not exist in the two-sided test by exemplifying the testing problem considered in this section. As we shown in Eq. (4.35). the rejection region of the UMP test is $x > c$, when the alternative hypothesis is $H_1 : \theta = \theta_1$ ($\theta_1 > 0$). On the other hand, $x < -c$ is the rejection region of the UMP test for the alternative hypothesis $H_1 : \theta = \theta_1$ ($\theta_1 < 0$). We want to find the test function for the two-sided test $H_0 : \theta = 0$ vs. $H_1 : \theta \neq 0$, which gives a UMP test for the alternative hypotheses, $\theta_1 > 0$ and $\theta_1 < 0$. simultaneously. However, the test functions of the UMP test are difference between these two case, $\theta_1 > 0$ and $\theta_1 < 0$. Therefore the UMP test does not exist in the two-sided test.

Consequently, the UMP test is an appropriate test for the one-sided test $H_0 : \theta = 0$ vs. $H_1 : \theta > 0$ (and also $H_0 : \theta \leq 0$ vs. $H_1 : \theta > 0$), and the UMPU test is appropriate test for the two-sided test $H_0 : \theta = 0$ vs. $H_1 : \theta \neq 0$ to consider the UMPU test. In other words, we should consider the UMPU test for the two-sided test rather than the UMP test or the likelihood-ratio test.

4.6 Summary of this chapter

In this chapter, we have seen the standard theory and process of the hypothesis testing method. In the hypothesis testing, we determine which hypothesis is true and obtain the information about a research objective. To proceed a good testing, we should prepare the appropriate test function corresponding to the testing problem, which enables us to reduce the probabilities of the two-type error, i.e., the misjudging. We consider the three types of testing problems and introduce the test function for each testing problem.

In the next chapter, we evaluate the hypothesis testing in the weak measurement for the testing problem to distinguish whether the interaction between a measured system and a measuring probe is presence or not.

[4]At $\theta \to \infty$, the powers converge 1.

Chapter 5

Weak-Value Amplification for Detection Problem

5.1 Preface

In this chapter, we consider the application of the hypothesis testing reviewed in the previous chapter to the interaction detection problem in the weak-value amplification. Here, we compare detection capabilities of the weak measurement and the conventional measurement. As stated below, the interaction detection problem should be treated as the two-sided test. For a fair comparison, we propose the test function which gives the UMPU test in each measurement case. By evaluating the probabilities of the type-1 and -2 errors and the statistical power, we can find that the weak measurement has an advantage to reduce the probability to miss the presence of the interaction when the weak value is outside the range of the eigenvalues. Additionally, we consider the proposed testing method under a typical noise. This chapter is based on Ref. [60].

5.2 Hypotheses and Test function for interaction detection

At first, we assume that the interaction Hamiltonian is given as the von Neumann type $\hat{H} = g\delta(t)\hat{A}\otimes\hat{p}$, the observable of the measured system is $\hat{A} = |+\rangle\langle+| - |-\rangle\langle-|$ for the two-state system, and the initial distribution of the measuring probe is given as the Gaussian

profile $\psi(x) = (2\pi\sigma^2)^{-1/4}\exp[-x^2/(4\sigma^2)]$ with the variance σ^2. Following the previous chapter, we assume that an obtained data is enough for statistical processing with the hypothesis testing method. More precisely, we consider the case that the number of the obtained data is infinitely large and the data loss by the failure of the post-selection is neglected.

To carry out the hypothesis testing method, we fix the two contradictory hypotheses, the null hypothesis and the alternative one. For the interaction detection, the following hypotheses are reasonable;

the null hypothesis H_0: the interaction is absent, i.e., $g = 0$,
the alternative hypothesis H_1: the interaction is present, i.e., $g \neq 0$.

Therefore we regard the interaction detection problem is the two-sided test. Here the type-1 and -2 errors mean as below[1]; the type-1 error represents that the interaction is really absent but we incorrectly guess the interaction exists, and the type-2 error shows that the interaction actually exists but we wrongly suppose there is no interaction, namely, we miss the presence of the interaction. The statistical power indicates the probability that we correctly judge the presence of the interaction when the interaction really exists. To calculate the probabilities of the errors and the power, we need the position distributions of the final probes after each measurement and the test function to determine from now on. Here we recapitulate the position distributions of the final probe given by the conventional measurement as

$$f_c(x|g) = \frac{1}{\sqrt{2\pi\sigma^2}}\left(|\langle +|i\rangle|^2 e^{-\frac{(x-g)^2}{2\sigma^2}} + |\langle -|i\rangle|^2 e^{-\frac{(x+g)^2}{2\sigma^2}}\right), \tag{5.1}$$

and the one observed in the weak measurement as

$$\begin{aligned} f_w(x|g) =& \frac{\mathcal{Z}^{-1}}{2\sqrt{2\pi\sigma^2}}\Big[(1 + |\langle\hat{A}\rangle_w|^2 + 2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x-g)^2}{2\sigma^2}} \\ &+(1 + |\langle\hat{A}\rangle_w|^2 - 2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x+g)^2}{2\sigma^2}} + 2(1 - |\langle\hat{A}\rangle_w|^2)e^{-\frac{x^2+g^2}{2\sigma^2}}\Big], \end{aligned} \tag{5.2}$$

$$\mathcal{Z} = 1 + |\langle\hat{A}\rangle_w|^2 + (1 - |\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}. \tag{5.3}$$

Next, on the basis of the Lemma 4.4.1 introduced in the previous chapter, we consider the test function $d(x)$ which gives the UMPU test for each measurement. Firstly, we need

[1]The type-1 and -2 errors equal to the false alarm and the false negative, respectively.

to prepare the appropriate test function as the candidate of the UMPU test, which does not contain the unknown coupling constant g. Here we aim to propose the test function from the physical intuition on the basis that the initial measuring probe distribution is the Gaussian, the variance of which is σ^2. If there is no interaction, i.e., the null hypothesis is true, we will obtain the measurement result x such that inside the initial probe fluctuation $|x| < \sigma$ rather than the outside $|x| > \sigma$. On the other hand, if the interaction is present, i.e., the alternative hypothesis is correct, we will observe the position data x such as outside the initial fluctuation $|x| > \sigma$ and the probability of $|x| < \sigma$ becomes comparatively small. This picture holds in both the measurements, the weak measurement and the conventional measurement. Therefore, we propose the following test function as the candidate of the UMPU test:

$$d(x) = \begin{cases} 0 & \text{if } |x|/\sigma < c, \\ r & \text{if } |x|/\sigma = c, \\ 1 & \text{if } |x|/\sigma > c, \end{cases} \tag{5.4}$$

where c is a critical point. We confirm that this test function gives the UMPU test or the UMP test in Sec. 5.4.

5.3 Comparison of the probabilities of the errors and the powers

Here, we calculate and compare the probabilities of the type-1 and -2 errors and the statistical power. First, we consider the probability of the type-1 error when the coupling constant is $g = 0$ and the measurement result is $|x| > c\sigma$ [Fig. 5.1(a)]. When the interaction is absent, the two distributions (5.1) and (5.2) correspond to the initial Gaussian distribution

$$f_c(x|g=0) = f_w(x|g=0) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}. \tag{5.5}$$

Therefore, the probability of the type-1 error of the each measurement are the same as

$$\Pr[\mathcal{E}_1] = \beta(0) = 1 - \mathrm{erf}\left[\frac{c}{\sqrt{2}}\right], \tag{5.6}$$

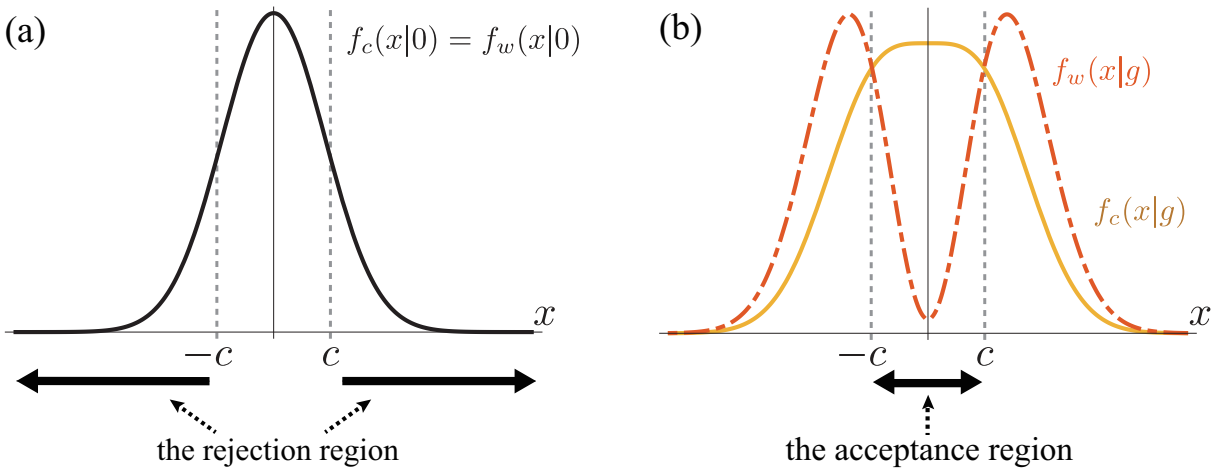


Figure 5.1: The schematic diagram of (a) the initial position distribution and (b) the final ones given by the conventional measurement and the weak measurement with the critical point. Also shown are (a) the rejection region and (b) the acceptance region to see the probabilities of the type-1 and -2 errors.

which is calculated by integrating the function (5.5) over the rejection region. Since $\Pr[\mathcal{E}_1]$ can be at any significance level by choosing c which is an arbitrary value, the test accommodates the standard strategy of the hypothesis testing.

Next, we see the probability of the type-2 error when the coupling constant is $g \neq 0$ and the measurement result is $|x| < c\sigma$ [Fig. 5.1(b)]. Since the interaction is present, the distributions (5.1) and (5.2) become different from each other. By integrating the function (5.1) over the acceptance region, the probability of the type-2 error given by the conventional measurement is calculated as

$$\begin{aligned}\Pr[\mathcal{E}_{2,c}] &= 1 - \beta_c(g) \\ &= 1 - \left(\int_{-\infty}^{-c\sigma} f_c(x|g)dx + \int_{c\sigma}^{\infty} f_c(x|g)dx \right) \\ &= \frac{1}{2}\left(\mathrm{erf}\left[\frac{c\sigma - g}{\sqrt{2\sigma^2}} \right] + \mathrm{erf}\left[\frac{c\sigma + g}{\sqrt{2\sigma^2}} \right] \right). \end{aligned} \tag{5.7}$$

Similarly the one in the case of the weak measurement is calculated as

$$\begin{aligned}
&\Pr[\mathcal{E}_{2,w}] \\
&= 1 - \beta_w(g) \\
&= 1 - \left(\int_{-\infty}^{-c\sigma} f_w(x|g)dx + \int_{c\sigma}^{\infty} f_w(x|g)dx \right) \\
&= \frac{1}{2\mathcal{Z}} \left[(1 + |\langle \hat{A} \rangle_w|^2) \left(\mathrm{erf} \left[\frac{c\sigma - g}{\sqrt{2\sigma^2}} \right] + \mathrm{erf} \left[\frac{c\sigma + g}{\sqrt{2\sigma^2}} \right] \right) + 2(1 - |\langle \hat{A} \rangle_w|^2) e^{-\frac{g^2}{2\sigma^2}} \mathrm{erf} \left[\frac{c}{\sqrt{2}} \right] \right].
\end{aligned} \tag{5.8}$$

Because the probability of the type-1 error is the same in both measurements, we compare the probabilities of the type-2 error to find which measurement is more advantageous. We can obtain and arrange the error-probability ratio $\Pr[\mathcal{E}_{2,w}]/\Pr[\mathcal{E}_{2,c}]$ as

$$\begin{aligned}
\frac{\Pr[\mathcal{E}_{2,w}]}{\Pr[\mathcal{E}_{2,c}]} - 1 &= \frac{1}{\mathcal{Z}} \left[(1 + |\langle \hat{A} \rangle_w|^2) + (1 - |\langle \hat{A} \rangle_w|^2) e^{-\frac{g^2}{2\sigma^2}} \left(\frac{2\mathrm{erf} \left[\frac{c}{\sqrt{2}} \right]}{\mathrm{erf} \left[\frac{c\sigma - g}{\sqrt{2\sigma^2}} \right] + \mathrm{erf} \left[\frac{c\sigma + g}{\sqrt{2\sigma^2}} \right]} \right) - \mathcal{Z} \right] \\
&= \mathcal{Z}^{-1} (1 - |\langle \hat{A} \rangle_w|^2) e^{-\frac{g^2}{2\sigma^2}} \left(\frac{2\mathrm{erf} \left[\frac{c}{\sqrt{2}} \right]}{\mathrm{erf} \left[\frac{c\sigma - g}{\sqrt{2\sigma^2}} \right] + \mathrm{erf} \left[\frac{c\sigma + g}{\sqrt{2\sigma^2}} \right]} - 1 \right).
\end{aligned} \tag{5.9}$$

Here we prove that the inequality

$$\frac{2\mathrm{erf} \left[\frac{c}{\sqrt{2}} \right]}{\mathrm{erf} \left[\frac{c\sigma - g}{\sqrt{2\sigma^2}} \right] + \mathrm{erf} \left[\frac{c\sigma + g}{\sqrt{2\sigma^2}} \right]} > 1 \tag{5.10}$$

always holds.

Proof. we can assume $g > 0$ without the loss of generality because the left-hand side of the inequality (5.10) is symmetric under the exchange $g \leftrightarrow -g$. In the case of $0 < g \leq c\sigma$, $e^{-(t - g\sqrt{2\sigma^2})^2} > e^{-t^2}$ holds when $t \leq c/\sqrt{2}$. Therefore, we obtain the inequality

$$\int_{\frac{c}{\sqrt{2}}}^{\frac{c\sigma + g}{\sqrt{2\sigma^2}}} e^{-(t - \frac{g}{\sqrt{2\sigma^2}})^2} dt > \int_{\frac{c}{\sqrt{2}}}^{\frac{c\sigma + g}{\sqrt{2\sigma^2}}} e^{-t^2} dt, \tag{5.11}$$

which becomes

$$\int_{\frac{c\sigma - g}{\sqrt{2\sigma^2}}}^{\frac{c}{\sqrt{2}}} e^{-t^2} dt > \int_{\frac{c}{\sqrt{2}}}^{\frac{c\sigma + g}{\sqrt{2\sigma^2}}} e^{-t^2} dt \Leftrightarrow \mathrm{erf} \left[\frac{c}{\sqrt{2}} \right] - \mathrm{erf} \left[\frac{c\sigma - g}{\sqrt{2\sigma^2}} \right] > \mathrm{erf} \left[\frac{c\sigma + g}{\sqrt{2\sigma^2}} \right] - \mathrm{erf} \left[\frac{c}{\sqrt{2}} \right]. \tag{5.12}$$

Thus, we have shown the inequality (5.10) for $0 < g \leq c\sigma$.

Next, we consider the case $g > c\sigma$. The inequality $e^{-(t-c/\sqrt{2})^2} > e^{-t^2}$ is given by $t \geq c/\sqrt{2}$, which provides

$$\int_{\frac{c}{\sqrt{2}}}^{\sqrt{2}c} e^{-(t-\frac{c}{\sqrt{2}})^2} dt > \int_{\frac{c}{\sqrt{2}}}^{\sqrt{2}c} e^{-t^2} dt \Leftrightarrow \int_{0}^{\frac{c}{\sqrt{2}}} e^{-t^2} dt > \int_{\frac{c}{\sqrt{2}}}^{\sqrt{2}c} e^{-t^2} dt$$
$$\Leftrightarrow \mathrm{erf}\left[\frac{c}{\sqrt{2}}\right] > \mathrm{erf}\left[\sqrt{2}c\right] - \mathrm{erf}\left[\frac{c}{\sqrt{2}}\right]. \tag{5.13}$$

In addition, from $e^{-(t-\sqrt{2}c)^2} > e^{-t^2}$ for $t \geq \sqrt{2}c$, we obtain

$$\int_{\sqrt{2}c}^{\frac{c\sigma+g}{\sqrt{2\sigma^2}}} e^{-\left(t-\sqrt{2}c\right)^2} dt > \int_{\sqrt{2}c}^{\frac{c\sigma+g}{\sqrt{2\sigma^2}}} e^{-t^2} dt \Leftrightarrow \int_{0}^{-\frac{c\sigma-g}{\sqrt{2\sigma^2}}} e^{-t^2} dt > \int_{\sqrt{2}c}^{\frac{c\sigma+g}{\sqrt{2\sigma^2}}} e^{-t^2} dt$$
$$\Leftrightarrow -\mathrm{erf}\left[\frac{c\sigma-g}{\sqrt{2\sigma^2}}\right] > \mathrm{erf}\left[\frac{c\sigma+g}{\sqrt{2\sigma^2}}\right] - \mathrm{erf}\left[\sqrt{2}c\right]. \tag{5.14}$$

Combining these inequalities, we acquire

$$\mathrm{erf}\left[\frac{c}{\sqrt{2}}\right] - \mathrm{erf}\left[\frac{c\sigma-g}{\sqrt{2\sigma^2}}\right] > \mathrm{erf}\left[\frac{c\sigma+g}{\sqrt{2\sigma^2}}\right] - \mathrm{erf}\left[\frac{c}{\sqrt{2}}\right] \tag{5.15}$$

for $g > c\sigma$. To summarize, we have proved the inequality (5.10). □

From Eq. (5.9) with the inequity (5.10), we can find that the inequality

$$\Pr[\mathcal{E}_{2,w}] \leq \Pr[\mathcal{E}_{2,c}] \tag{5.16}$$
$$\Leftrightarrow \beta_w(g) \geq \beta_c(g) \tag{5.17}$$

holds for the weak value satisfying

$$|\langle \hat{A} \rangle_w| \geq 1, \tag{5.18}$$

where "1" in the right-hand side is derived from the eigenvalue of the observable $\hat{A}$. Hence, the probability of the type-2 error can be reduced by the appropriate postselection. We note that the ratio $\Pr[\mathcal{E}_{2,w}]/\Pr[\mathcal{E}_{2,c}]$ is a monotonically decreasing function with respect

to $|\langle\hat{A}\rangle_w|^2$ because its derivative is negative,

$$\frac{\partial \Pr[\mathcal{E}_{2,w}]/\Pr[\mathcal{E}_{2,c}]}{\partial |\langle\hat{A}\rangle_w|^2} = -2\mathcal{Z}^2 \left(\frac{2\mathrm{erf}\left[\frac{c}{\sqrt{2}}\right]}{\mathrm{erf}\left[\frac{c\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c\sigma + g}{\sqrt{2\sigma^2}}\right]} - 1 \right) e^{-\frac{g^2}{2\sigma^2}} \leq 0. \tag{5.19}$$

Also the ratio of the powers β_w/β_c is a monotonically increasing function with regard to $|\langle\hat{A}\rangle_w|^2$.

We remark that the merit of the WVA is originally supposed to that amplifying the shift of the expectation value by weak value larger than the eigenvalue. Here we additionally give the implication of the WVA that reduces the possibility of missing the presence of the interaction, which is mathematically well-grounded. We suppose that this is the one of the essences of the WVA.

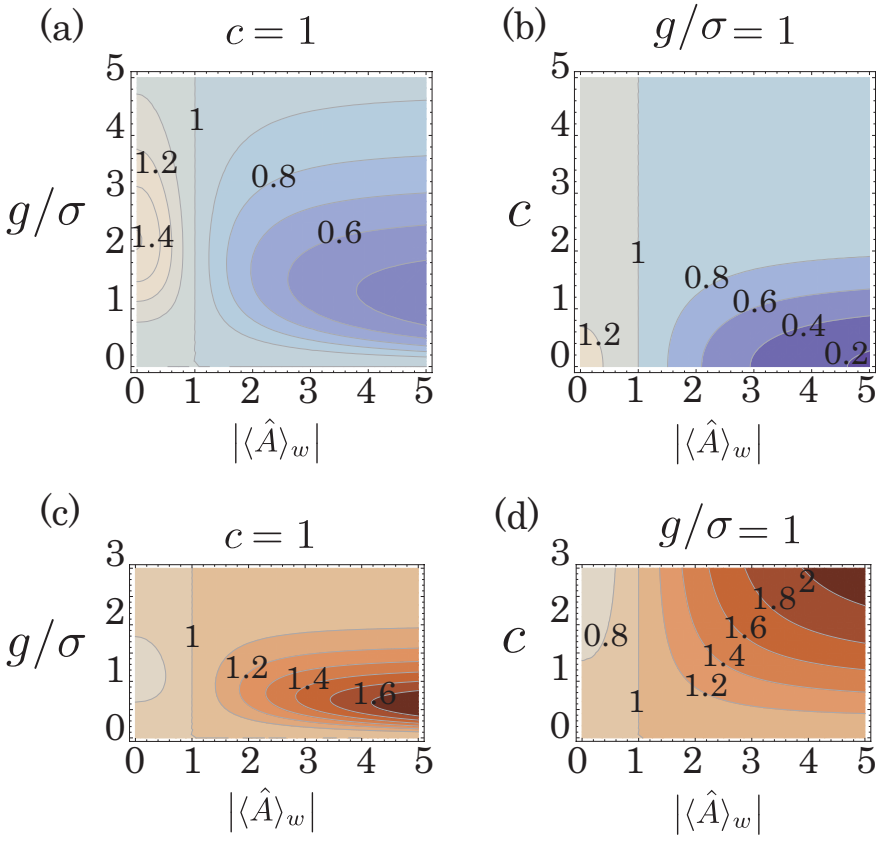


Figure 5.2: The contour plots of (a), (b) the ratios of the two probabilities of the type-2 error $\Pr[\mathcal{E}_{2,w}]/\Pr[\mathcal{E}_{2,c}]$ and (c), (d) the ratio of the two powers $\beta_w(g)/\beta_c(g)$. The horizontal axis of these plots shows the absolute value of the weak value $|\langle\hat{A}\rangle_w|$. The vertical axis indicate the interaction strength divided by the initial fluctuation g/σ in (a) and (c), and the critical point c in (b) and (d). The darker blue (red) indicates the small (large) values.

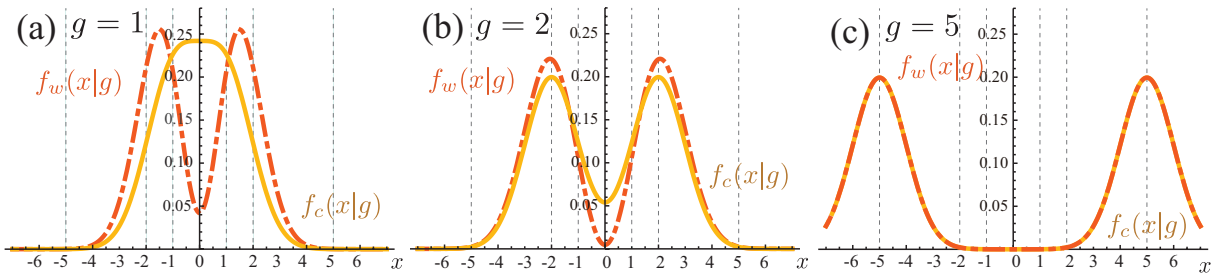


Figure 5.3: The transition of the position distributions given by the conventional measurement $f_c(x|g)$ (the orange solid curve) and the weak measurement $f_w(x|g)$ (the red dashed curve) to the interaction strength. In (c), the two plots are almost overlapped. The parameters are fixed as $\mathrm{Re}\langle\hat{A}\rangle_w = 0$, $|\langle\hat{A}\rangle_w| = 5$, $|\langle+|i\rangle|^2 = |\langle-|i\rangle|^2 = 1/2$, and $\sigma = 1$.

We plot the probability ratio $\Pr[\mathcal{E}_{2,w}]/\Pr[\mathcal{E}_{2,c}]$ and the power ratio $\beta_w(g)/\beta_c(g)$ for the parameters $|\langle\hat{A}\rangle_w|$, g/σ and c in Fig. 5.2, which indicate that the inequalities (5.16) and (5.17) hold for $|\langle\hat{A}\rangle_w| \geq 1$. We can see that the WVA works well when g/σ and c are relatively small, but are not zero. On the other hand, if the g/σ is large, the weak measurement is as helpful as the conventional measurement for the detection of the interaction. This property comes from the difference of the distribution functions $f_c(x|g)$ and $f_w(x|g)$, which can be explained in Fig. 5.3 that shows the transition of the position distributions $f_c(x|g)$ and $f_w(x|g)$ to the interaction strength.

There is a big difference between $f_c(x|g)$ and $f_w(x|g)$ in the small $|x|$ region for the small g as shown in Fig. 5.3 (a). Meanwhile $f_c(x|g)$ and $f_w(x|g)$ are almost same for a large g as we can see from Fig. 5.3 (b) and (c). As we know that the probability of the type-2 error is given by the integration over the acceptance region, i.e., the interval $[-c\sigma, c\sigma]$. Since $f_w(x|g)$ is smaller than $f_c(x|g)$ in the central region of x for a fixed small g, the ratio $\Pr[\mathcal{E}_{2,w}]/\Pr[\mathcal{E}_{2,c}]$ becomes small, if we properly choose the critical point c. In contrast, the ratio of powers $\beta_w(g)/\beta_c(g)$ becomes large.

5.4 Confirmation that the test is UMPU or UMP

5.4.1 Proof of the UMPU test

We prove that the test function (5.4) provides the UMPU test for the each measurement according to the Lemma 4.4.1. Firstly we see Eqs. (4.22), (4.23), and (4.24). Here we

check Eq. (4.22). In the weak measurement, we can obtain the equations

$$\begin{aligned}\partial_g\beta_w(g) &= \partial_g(1-\Pr[\mathcal{E}_{2,w}]) \\ &= \frac{\mathcal{Z}^{-1}}{\sqrt{2\pi\sigma^2}}(1+|\langle\hat{A}\rangle_w|^2)\big(e^{-\frac{(c\sigma-g)^2}{2\sigma^2}}-e^{-\frac{(c\sigma+g)^2}{2\sigma^2}}\big) \\ &\quad+\frac{Z^{-1}g}{\sigma^2}(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}\left(\mathrm{erf}\left[\frac{c}{\sqrt{2}}\right]-\Pr[\mathcal{E}_{2,ps}]\right), \end{aligned} \tag{5.20}$$

and

$$\begin{aligned}\int d(x)\frac{\partial}{\partial g}f_w(x|g)dx &= \frac{\mathcal{Z}^{-1}}{2\sqrt{2\pi\sigma^2}}\int d(x)\frac{\partial}{\partial x}\Big(-(1+|\langle\hat{A}\rangle_w|^2+2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x-g)^2}{2\sigma^2}} \\ &\qquad +(1+|\langle\hat{A}\rangle_w|^2-\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x+g)^2}{2\sigma^2}}\Big)dx \\ &\quad-\frac{\mathcal{Z}^{-1}g}{2\sigma^2}(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}\int d(x)\frac{\partial}{\partial x}\mathrm{erf}\left[\frac{x}{\sqrt{2}\sigma}\right]dx \\ &\quad-\mathcal{Z}^{-1}\frac{\partial Z}{\partial g}\int d(x)f_w(x|g)dx \\ &= \frac{\mathcal{Z}^{-1}}{\sqrt{2\pi\sigma^2}}(1+|\langle\hat{A}\rangle_w|^2)\big(e^{-\frac{(c\sigma-g)^2}{2\sigma^2}}-e^{-\frac{(c\sigma+g)^2}{2\sigma^2}}\big) \\ &\quad+\frac{Z^{-1}g}{\sigma^2}(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}\left(\mathrm{erf}\left[\frac{c}{\sqrt{2}}\right]-\Pr[\mathcal{E}_{2,ps}]\right), \end{aligned} \tag{5.21}$$

which satisfy Eq. (4.22). Similarly, in the conventional measurement, we obtain the equations

$$\partial_g\beta_c(g) = \partial_g(1-\Pr[\mathcal{E}_{2,c}]) = \frac{1}{\sqrt{2\pi\sigma^2}}\big(e^{-\frac{(c\sigma-g)^2}{2\sigma^2}}-e^{-\frac{(c\sigma+g)^2}{2\sigma^2}}\big) \tag{5.22}$$

and

$$\begin{aligned}\int d(x)\frac{\partial}{\partial g}f_c(x|g)dx &= \frac{1}{\sqrt{2\pi\sigma^2}}\int d(x)\frac{\partial}{\partial x}\left(-|\langle+|i\rangle|^2e^{-\frac{(x-g)^2}{2\sigma^2}}+|\langle-|i\rangle|^2e^{-\frac{(x+g)^2}{2\sigma^2}}\right)dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\big(e^{-\frac{(c\sigma-g)^2}{2\sigma^2}}-e^{-\frac{(c\sigma+g)^2}{2\sigma^2}}\big), \end{aligned} \tag{5.23}$$

which also satisfy Eq. (4.22). According to Eq. (5.6), the probability of the type-1 error can be the significance level α by choosing a proper value for the critical point c, which means that Eq. (4.23) is satisfied in the each measurement. We can also see that the test function $d(x)$ satisfies Eq. (4.24) in the each measurement from Eqs. (5.20) and (5.20).

Next, we consider Eq. (4.26). In the case of the weak measurement which gives the distribution $f_w(x|g)$, the equation becomes

$$\frac{\mathcal{F}_w(x)}{f_w(x|0)} = \mathcal{G}_w(x) - c_{1,w} - c_{2,w}\frac{\mathrm{Re}\langle\hat{A}\rangle_w}{\sigma^2}x, \tag{5.24}$$

$$\mathcal{G}_w(x) := \frac{e^{-\frac{g^2}{2\sigma^2}}}{2\mathcal{Z}}\Big[(1+|\langle\hat{A}\rangle_w|^2+2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{xg}{\sigma^2}} + (1+|\langle\hat{A}\rangle_w|^2-2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{xg}{\sigma}} + 2(1-|\langle\hat{A}\rangle_w|^2)\Big]. \tag{5.25}$$

When $c_{1,w}$ and $c_{2,w}$ are

$$c_{1,w} = \frac{\mathcal{G}_w(c\sigma)+\mathcal{G}_w(-c\sigma)}{2} = \frac{e^{-\frac{g^2}{2\sigma^2}}}{\mathcal{Z}}\left[2(1+|\langle\hat{A}\rangle_w|^2)\cosh\left[\frac{cg}{\sigma}\right] + (1-|\langle\hat{A}\rangle_w|^2)\right], \tag{5.26}$$

$$c_{2,w} = \sigma\frac{\mathcal{G}_w(c\sigma)-\mathcal{G}_w(-c\sigma)}{2\mathrm{Re}\langle\hat{A}\rangle_w c} = \frac{2\sigma e^{-\frac{g^2}{2\sigma^2}}\sinh\left[\frac{cg}{\sigma^2}\right]}{c\mathcal{Z}}, \tag{5.27}$$

$x = \pm c\sigma$ become the solutions of $\mathcal{F}_w(x)/f_w(x|0) = 0$ for $\mathrm{Re}\langle\hat{A}\rangle_w \neq 0$. Since $\mathcal{G}_w(x)$ is a convex function, the following relation holds;

$$d(x) = \begin{cases} 0 & \text{if } \mathcal{F}_w(x) < 0 \Leftrightarrow |x|/\sigma < c, \\ r & \text{if } \mathcal{F}_w(x) = 0 \Leftrightarrow |x|/\sigma = c, \\ 1 & \text{if } \mathcal{F}_w(x) > 0 \Leftrightarrow |x|/\sigma > c. \end{cases} \tag{5.28}$$

We can see that the test function (5.4) gives the UMPU test for the weak measurement. Even if $\mathrm{Re}\langle\hat{A}\rangle_w = 0$, the argument remains valid.

Next, we consider the case of the conventional measurement. With the distribution $f_c(x|g)$, the equation becomes

$$\frac{\mathcal{F}_c(x)}{f_c(x|0)} = \mathcal{G}_c(x) - c_{1,c} - c_{2,c}\frac{|\langle+|i\rangle|^2-|\langle-|i\rangle|^2}{\sigma^2}x, \tag{5.29}$$

$$\mathcal{G}_c(x) := e^{-\frac{g^2}{2\sigma^2}}\left(|\langle+|i\rangle|^2 e^{\frac{xg}{\sigma^2}} + |\langle-|i\rangle|^2 e^{-\frac{xg}{\sigma^2}}\right). \tag{5.30}$$

When $c_{1,c}$ and $c_{2,c}$ are

$$c_{1,c} = \frac{\mathcal{G}_c(c\sigma)+\mathcal{G}_c(-c\sigma)}{2} = e^{-\frac{g^2}{2\sigma^2}}\cosh\left[\frac{cg}{\sigma}\right], \tag{5.31}$$

$$c_{2,c} = \frac{\mathcal{G}_c(c\sigma)-\mathcal{G}_c(-c\sigma)}{2(|\langle+|i\rangle|^2-|\langle-|i\rangle|^2)c} = \frac{\sigma e^{-\frac{g^2}{2\sigma^2}}}{c}\sinh\left[\frac{cg}{\sigma}\right], \tag{5.32}$$

$x = \pm c\sigma$ become the solutions of $\mathcal{F}_c(x)/f_c(x|0) = 0$. Here, we have assumed $|\langle +|i\rangle|^2 \neq \langle -|i\rangle|^2$. Even if $|\langle +|i\rangle|^2 = \langle -|i\rangle|^2$, the discussion works. Because $\mathcal{G}_c(x)$ is a convex function, we obtain the following relation;

$$d(x) = \begin{cases} 0 & \text{if } \mathcal{F}_c(x) < 0 \Leftrightarrow |x|/\sigma < c, \\ r & \text{if } \mathcal{F}_c(x) = 0 \Leftrightarrow |x|/\sigma = c, \\ 1 & \text{if } \mathcal{F}_c(x) > 0 \Leftrightarrow |x|/\sigma > c. \end{cases} \tag{5.33}$$

We have shown that the test function (5.4) provides the UMPU test in both measurements.

5.4.2 The case of the UMP test

We can show that the our test function (5.4) gives the UMP test in the case of $\mathrm{Re}\langle \hat{A}\rangle_w = 0$ for the weak measurement and in the case of $|\langle +|i\rangle|^2 = |\langle -|i\rangle|^2$ for the conventional measurement with the Theorem 4.3.1 presented in Sec. 4.3. With these conditions, the final distributions (5.1) and (5.2) become even function of g. We cannot read out the sign of g from the final position distributions. Thus, it is practically allowed to assume the sign of g is positive without loss of generality. Because of it, we can interpret that the hypotheses becomes $H_0 : g = 0$ and $H_1 : g > 0$ which is the one-sided test. Therefore, the Theorem can be applied to the derivation of the UMP test for the current interaction detection problem with the particular conditions.

From Eqs. (5.1) and (5.2), the likelihood-ratio in each measurement case is given as

$$\frac{f_w(x|g)}{f_w(x|g=0)} = \frac{1 - (\mathrm{Im}\langle \hat{A}\rangle_w)^2 + \left[1 + (\mathrm{Im}\langle \hat{A}\rangle_w)^2\right] \cosh\left[\frac{xg}{\sigma^2}\right]}{1 - (\mathrm{Im}\langle \hat{A}\rangle_w)^2 + \left[1 + (\mathrm{Im}\langle \hat{A}\rangle_w)^2\right] e^{\frac{g^2}{2\sigma^2}}}, \tag{5.34}$$

$$\frac{f_c(x|g)}{f_c(x|g=0)} = e^{-\frac{g^2}{2\sigma^2}} \cosh\left[\frac{xg}{\sigma^2}\right], \tag{5.35}$$

respectively. Here we have used $\mathrm{Re}\langle \hat{A}\rangle_w = 0$ and $|\langle +|i\rangle|^2 = |\langle -|i\rangle|^2$. Because cosh is an even convex function, the following holds;

$$\cosh\left[\frac{gx}{\sigma^2}\right] = \cosh\left[\frac{g}{\sigma}\frac{|x|}{\sigma}\right]. \tag{5.36}$$

Since the interaction strength g to be measured is the unknown parameter, the statistics $T(x)$ does not have to contain the parameter. It is valid that the statistics is fixed as $T(x) = |x|/\sigma$. Because the statistics $T(x)$ is a monotonically increasing function, we find

that $T(x) = |x|/\sigma$ gives the UMP test for the weak measurement with $\mathrm{Re}\langle\hat{A}\rangle_w = 0$ and for the conventional measurement with $|\langle +|i\rangle|^2 = |\langle -|i\rangle|^2$.

5.5 Testing under additive white Gaussian noise

In this section, we discuss the testing method considered under the noise. We derive the distributions of the measurement result x as (5.1) and (5.2) which can be obtained in ideal experiments. Practically, the noise influence is inevitable and we should take it into account. Here, we suppose the noise y which is added to x by passing through the device circuit. The noise y takes the Gaussian probability $N(y) := e^{-y^2/2s^2}/\sqrt{2\pi s^2}$, where s is an arbitrary fluctuation. This noise model is known as an additive white Gaussian noise. The noise occurs by the thermal noise or the shot noise in an electric device.

To discuss the noise influence, we derive the distribution $z(:= x+y)$ from the moment-generating function of x and y. The moment-generating function of the Gaussian function $N(y)$ is given as

$$\mathrm{E}[e^{ty}] = \int e^{ty} N(y) dy = e^{\frac{s^2}{2}t^2}, \tag{5.37}$$

where the $\mathrm{E}[\cdot]$ represents the expectation value with the Gaussian function $N(y)$. Also we use the notation $\mathrm{E}_{w(c)}[\cdot]$ represents the expectation value with the distribution function $f_{w(c)}(x|g)$. The moment-generating function of the x distribution given in the weak measurement is

$$\begin{aligned}\mathrm{E}_w[e^{tx}] =& \int e^{tx} f_w(x|g) dx \\ =& \frac{e^{\frac{\sigma^2}{2}t^2}}{2s\mathcal{Z}} \Big[(1 + |\langle\hat{A}\rangle_w|^2 + 2\mathrm{Re}\langle\hat{A}\rangle_w) e^{gt} \\ & + (1 + |\langle\hat{A}\rangle_w|^2 - 2\mathrm{Re}\langle\hat{A}\rangle_w) e^{-gt} + 2(1 - |\langle\hat{A}\rangle_w|^2) e^{-\frac{g^2}{2\sigma^2}} \Big]. \end{aligned} \tag{5.38}$$

Because the random variable x and y are independent, the moment-generating function of $z = x + y$ fulfills

$$\begin{aligned}\mathrm{E}_w[e^{tz}] =& \mathrm{E}_w[e^{tx}]\mathrm{E}[e^{ty}] \\ =& \frac{e^{\frac{\sigma^2+s^2}{2}t^2}}{2\mathcal{Z}}\Big[(1+|\langle\hat{A}\rangle_w|^2+2\mathrm{Re}\langle\hat{A}\rangle_w)e^{gt} \\ &+(1+|\langle\hat{A}\rangle_w|^2-2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-gt}+2(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}\Big]. \end{aligned} \tag{5.39}$$

Because of the linearity of the expectation value, the distribution of z becomes

$$\begin{aligned}& f_w(z|g) \\ =& \frac{\mathcal{Z}^{-1}}{2\sqrt{2\pi(\sigma^2+s^2)}}\Big[(1+|\langle\hat{A}\rangle_w|^2+2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x-g)^2}{2(\sigma^2+s^2)}} \\ &+(1+|\langle\hat{A}\rangle_w|^2-2\mathrm{Re}\langle\hat{A}\rangle_w)e^{-\frac{(x+g)^2}{2(\sigma^2+s^2)}}+2(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}e^{-\frac{x^2}{2(\sigma^2+s^2)}}\Big]. \end{aligned} \tag{5.40}$$

Likewise, we can obtain the moment-generating function of the x distribution given in the conventional measurement as

$$\mathrm{E}_c[e^{tx}] = e^{\frac{\sigma^2}{2}t^2}(|\langle+|i\rangle|^2e^{gt}+|\langle-|i\rangle|^2e^{-gt}), \tag{5.41}$$

and the distribution function of z as

$$f_c(z|g) = \frac{1}{\sqrt{2\pi(\sigma^2+s^2)}}(|\langle+|i\rangle|^2e^{-\frac{(z-g)^2}{2(\sigma^2+s^2)}}+|\langle-|i\rangle|^2e^{-\frac{(z+g)^2}{2(\sigma^2+s^2)}}), \tag{5.42}$$

We have derived the distribution functions with the Gaussian noise (5.40) and (5.42), from which we calculate the probabilities of the two errors.

The probability of the type-1 error given in the absence of the interaction are the same in the two measurements as

$$\Pr[\mathcal{E}_{1,c}] = \Pr[\mathcal{E}_{1,w}] = 1-\mathrm{erf}\left[\frac{c\sigma}{\sqrt{2(\sigma^2+s^2)}}\right]. \tag{5.43}$$

The probability of the type-2 error given in the presence of the interaction is calculated as

$$\Pr[\mathcal{E}_{2,w}] = \frac{1}{2\mathcal{Z}}\left[(1+|\langle\hat{A}\rangle_w|^2)\left(\mathrm{erf}\left[\frac{c\sigma-g}{\sqrt{2(\sigma^2+s^2)}}\right]+\mathrm{erf}\left[\frac{c\sigma+g}{\sqrt{2(\sigma^2+s^2)}}\right]\right)\right.$$
$$\left.+(1-|\langle\hat{A}\rangle_w|^2)e^{-\frac{g^2}{2\sigma^2}}\mathrm{erf}\left[\frac{c\sigma}{\sqrt{2(\sigma^2+s^2)}}\right]\right] \tag{5.44}$$

for the weak measurement, while the one for the conventional measurement is

$$\Pr[\mathcal{E}_{2,c}] = \frac{1}{2}\left(\mathrm{erf}\left[\frac{c\sigma+g}{\sqrt{2(\sigma^2+s^2)}}\right]+\mathrm{erf}\left[\frac{c\sigma-g}{\sqrt{2(\sigma^2+s^2)}}\right]\right). \tag{5.45}$$

These probabilities with noise is almost the same as the ones considered in the previous chapter. We can obtain the same inequality as Eq. (5.16) which holds for $|\langle\hat{A}\rangle_w|^2 \geq 1$ by comparing Eqs. (5.45) and (5.45). Therefore, our conclusion is valid even if there is the additive white Gaussian noise.

5.6 Summary of this chapter

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. As we introduced in Sec. 2.4, the amplification effect of the weak measurement cannot enhance the estimation capability even if we have the infinitely large number of data. In this chapter, we find the specific case that the amplification effect becomes significant for the infinitely large number of data, which is the interaction detection. We have proposed the test function, which represents the determination criteria, where we suppose that the interaction is absent when the obtained data is inside the initial fluctuation of the probe and the interaction is present when the obtained data is outside the initial fluctuation of the probe. Consequently, if $\left|\langle\hat{A}\rangle_w\right| \geq 1$, the WVA has the advantage for the reduction of the type 2 error while keeping the type 1 error fixed, regardless of the value of the coupling constant g, the initial fluctuation σ and the critical point c. Namely, the weak measurement can decrease the probability of missing the presence of the interaction more than the conventional measurement when the weak value is outside the normal range of the eigenvalues. Furthermore, in both measurements, the false alarm rate is the same. We have also shown that our conclusion holds even if the additive white Gaussian noise is

present. For our discussion, we assume that the measured system is the two-state system, the initial wave function of the measuring probe is of the Gaussian probe, the number of the obtained data is infinitely large, and further the data loss by the failure of the postselection is neglected.

Our proposed test function gives the UMPU test for the interaction detection problem, which should be treated as the two-sided test. In the specific case such that the detection problem essentially behaves as a one-sided test, the proposed test can be regarded as a UMP test. We remark that, 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.

Note that we give further discussion about the hypothesis testing for the weak measurement in Appendix D. In Appendix D.2, we consider the case that the range of null hypothesis is a small interval such as $|g| \leq \varepsilon$ for small ε. This discussion is motivated by Refs. [56, 57], the author of which mentions that one might care a point null hypothesis such as $g = 0$. That is to say, with such a point null hypothesis, we will judge that the interaction is present although the interaction strength is very small such as $g \approx 10^{-10}$. Because of that, we suppose that the point null hypothesis does not always provide a reasonable determination. There we see the general tendency that the weak measurement becomes advantageous when the error of missing the presence of the interaction is more serious than the misdetection. Additionally, one may want to discuss the case including the data loss by the failure of the postselection. In Appendix D.3, we alternatively propose the makeshift decision function in order to discuss such a case, which has a few defects in its physical interpretation. Using the Lagrange multiplier method, we conclude that the measurement without the postselection gives the stationary solution to the Lagrangian.

In next chapter, we apply the proposed testing method for the actual optical experiment.

Chapter 6

Experimental description of the testing method

6.1 Preface

In this chapter, we see the testing method with the weak-value amplification through the experimental setup of the weak measurement. Among several experiments, we pick up the classic experiment using a single birefringent crystal and two polarizers demonstrated by Ritchie *et al.* [20]. This experiment was originally designed for measuring the weak value. Here we regard it as the experimental setup for the testing problem to distinguish whether the crystal used is birefringent or not. We conclude that the appropriate angle of the polarizers gives the advantage of the weak measurement. This chapter is based on the work [61].

6.2 Weak measurement experiment using a birefringent crystal

In this section, we see the experimental setup and its result demonstrated by Ritchie *et al.* [20]. Figure 6.1 (a) displays the optical setup, and the polarization and the position distribution of the probe in each stage.

In this weak measurement experiment, the position x of the laser beam is regarded as the measuring probe, the initial distribution of which is of the Gaussian shape (2.1) as

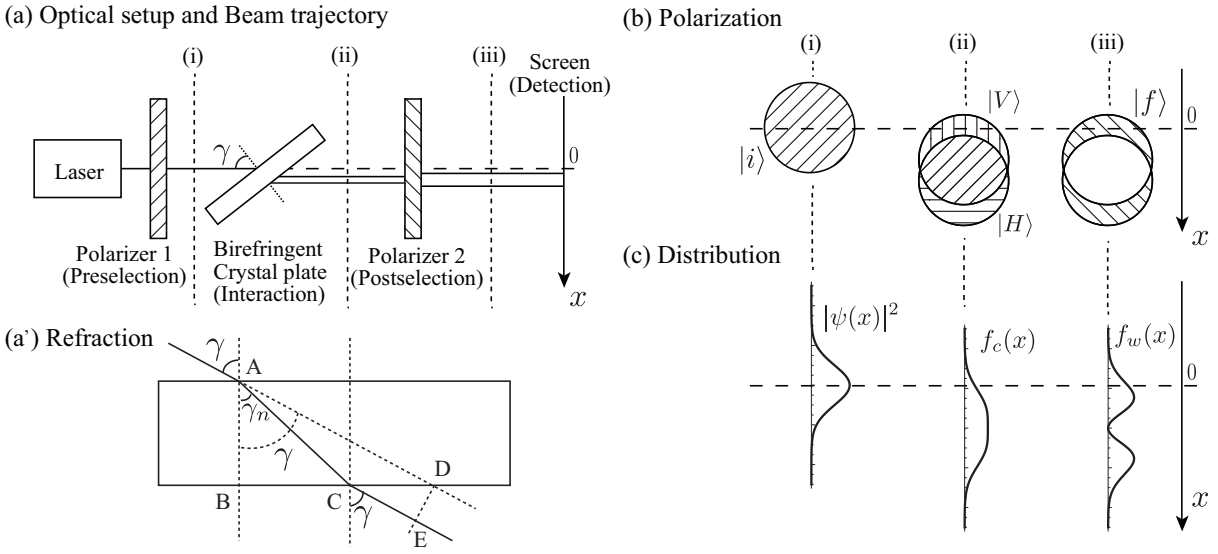


Figure 6.1: (a) Sketch of the optical setup. The tilted birefringent crystal plates is set between the two polarizers, the angles of which are tuned almost orthogonal. On the screen, we observe the beam position y and we decide whether or not the crystal is birefringent. (a') shows the refraction by the crystal plate. Also shown are (b) the beam polarization and (c) the probe position distribution, in each stage. For the illustration, the angles of the polarizers are taken as $\theta_i = \pi/4$ and $\theta_f = 3\pi/4$.

shown in Fig. 6.1(c)(i). The beam waist used in the experiment is $\sigma = 55\,\mu$m. The tuning of the two polarizers corresponds to the pre- and postselection of the initial and final states of the measured system, which are described as

$$|i\rangle = \cos\theta_i|H\rangle + \sin\theta_i|V\rangle, \tag{6.1}$$

$$|f\rangle = \cos\theta_f|H\rangle + \sin\theta_f|V\rangle, \tag{6.2}$$

where $|H\rangle$ and $|V\rangle$ are the horizontal and vertical polarization states, respectively, and θ_i and θ_f are the angles of the first and second polarizers, respectively [Fig. 6.1(c)(i) and (iii)]. A birefringent crystal is fixed between the two polarizers, which gives the two different refraction factors depending on the polarization of the injected beam. Because of the different refraction factors, the injected beam is spatially separated into the two beams with the different polarization [Fig. 6.1(b)(ii)], one of which is called an ordinary ray and the other is an extraordinary ray. The birefringent crystal induces the correlation between the measuring probe and the measured system. The refraction, which is the translation of injected beam, is given by the von Neumann type Hamiltonian (2.3). In this experiment, due to the two different refraction factors, the observable $\hat{A}$ of the measured system

becomes

$$\hat{A} = \lambda_H |H\rangle\langle H| + \lambda_V |V\rangle\langle V|, \tag{6.3}$$

where λ_H and λ_V are the eigenvalues of the polarization states $|H\rangle$ and $|V\rangle$, respectively. The birefringent crystal used in this experiment is a quartz plate, the refraction factors of which are $n_e = 1.551\,65$ for the extraordinary ray and $n_0 = 1.542\,61$ for the ordinary ray when the wavelength of the injected laser is 633 nm as quoted in Ref. [20]. Now the weak value is given as

$$\langle \hat{A} \rangle_w = \left(\frac{\lambda_H - \lambda_V}{2} \right) \frac{\cos(\theta_i + \theta_f)}{\cos(\theta_i - \theta_f)} + \frac{\lambda_H + \lambda_V}{2}. \tag{6.4}$$

As we can see, the weak value can be arbitrarily large when the postselected state is almost orthogonal to the preselection state.

After refraction by the birefringent crystal, the position distribution function of the probe becomes

$$f_c(x) = |\langle x|e^{-ig\hat{A}\otimes\hat{p}}|\psi\rangle|i\rangle|^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \left(\cos^2\theta_i e^{-\frac{(x-g\lambda_H)}{2\sigma^2}} + \sin^2\theta_i e^{-\frac{(x-g\lambda_V)}{2\sigma^2}} \right), \tag{6.5}$$

which is obtained as the output of the conventional measurement [Fig. 6.1(c)(ii)]. This distribution has two Gaussian functions. Here we evaluate the refraction shifts $g\lambda_H$ and λ_V. Figure 6.1(a') shows the refraction by the crystal plate. In this figure, λ_n is the refraction angle for the refraction factor n, the length of DE is the refraction shift. Using Snell's law,we can obtain $\sin\gamma = n\sin\gamma_n$. According to Fig. 6.1(a'), the length of DE is calculated as

$$\mathrm{DE} = \mathrm{CD}\sin(\pi/2 - \gamma) = [\mathrm{BD} - \mathrm{BC}]\cos\gamma = \mathrm{AB}[\tan\gamma - \tan\gamma_n]\cos\gamma = d\frac{\sin(\gamma - \gamma_n)}{\cos\gamma_n}. \tag{6.6}$$

The thickness of the crystal plate is $d = 331$ nm and the injected angle is $\gamma \approx 30°$ according to Ref. [20]. The refraction shifts are

$$g\lambda_H = d\frac{\sin(\gamma - \gamma_e)}{\cos\gamma_e} \approx 67.92\,\mu\mathrm{m}, \tag{6.7}$$

$$g\lambda_V = d\frac{\sin(\gamma - \gamma_o)}{\cos\gamma_o} \approx 67.28\,\mu\mathrm{m}, \tag{6.8}$$

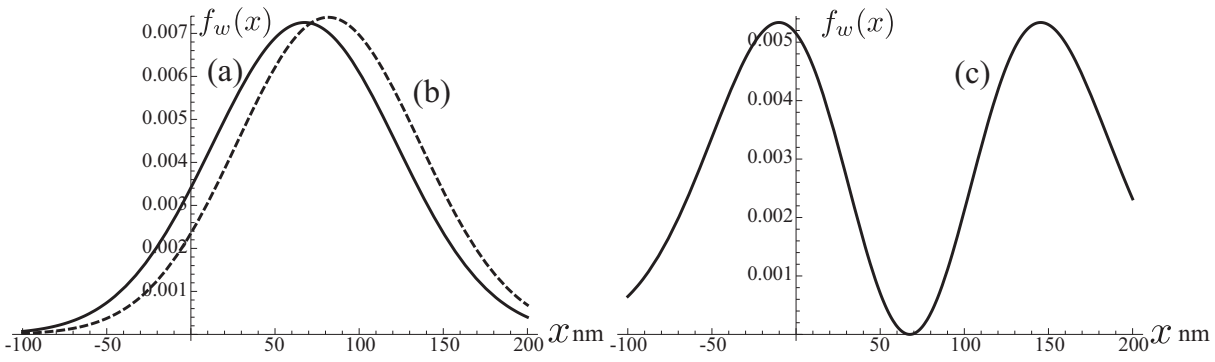


Figure 6.2: Plots of the position distribution $f_w(x)$ obtained by the weak measurement in three cases:(a) $\theta_i = \pi/4$ and $\theta_f = \pi/4$, (b) $\theta_i = \pi/4$ and $\theta_f = 3\pi/4 + 2.2 \times 10^{-}2$, (c) $\theta_i = \pi/4$ and $\theta_f = 3\pi/4$. The other parameters are fixed as $g\lambda_H \approx 67.92\mu$m, $g\lambda_V \approx 67.28\mu$m, and $\sigma = 55\mu$m.

where $\sin\gamma_e = \sin\gamma/n_e$ and $\sin\gamma_o = \sin\gamma/n_o$. Here we have assumed that the horizontal and vertical polarized beams are the extraordinary and the ordinary rays, respectively.

The position distribution is altered from $f_c(x)$ by postselection of the second polarizer, which is given as

$$\begin{aligned} f_w(x) &= \frac{|\langle x|\langle f|e^{-ig\hat{A}\otimes\hat{p}}|i\rangle|\psi\rangle|^2}{|\langle f|e^{-ig\hat{A}\otimes\hat{p}}|i\rangle|\psi\rangle|^2} \\ &= \frac{\cos^2\theta_i\cos^2\theta_f e^{-\frac{(x-g\lambda_H)^2}{2\sigma^2}} + \sin^2\theta_i\sin^2\theta_f e^{-\frac{(x-g\lambda_V)^2}{2\sigma^2}} + \frac{1}{2}\sin 2\theta_i\sin 2\theta_f e^{-\frac{1}{2\sigma^2}\left\{x-g\left(\frac{\lambda_H+\lambda_V}{2}\right)^2\right\}-\frac{g^2}{2\sigma^2}\left(\frac{\lambda_H-\lambda_V}{2}\right)^2}}{\sqrt{2\pi\sigma^2}\left(\cos^2\theta_i\cos^2\theta_f + \sin^2\theta_i\sin^2\theta_f + \frac{1}{2}\sin 2\theta_i\sin 2\theta_f e^{-\frac{g^2}{2\sigma^2}\left(\frac{\lambda_H-\lambda_V}{2}\right)^2}\right)}. \end{aligned} \quad (6.9)$$

This position probability distribution is obtained by the weak measurement. The third term of the numerator is generated by postselection.

Figure 6.2 depicts $f_w(x)$ in the three cases which the experimental results shown in Ref. [20]. In this experiment, Ritchie *et al.* tuned the angels θ_i and θ_f in the three cases: (a) $\theta_i = \pi/4$ and $\theta_f = \pi/4$ which can be almost regarded as the case of the conventional measurement[1], (b) $\theta_i = \pi/4$ and $\theta_f = 3\pi/4+2.2\times10^{-2}$, i.e., the case of weak measurement, (c) $\theta_i = \pi/4$ and $\theta_f = 3\pi/4$, i.e., the weak measurement with postselection which is completely orthogonal to the preselected state. We can see that the peak of distribution in (b) is shifted larger than the one in (a). In the case (b), because the angles of polarizers meet the AAV approximation $g|\langle\hat{A}\rangle_w| \ll 1$, the distribution is shifted while keeping the

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

single Gaussian. On the other hand, in the case (c), the AAV approximation brakes down so that the final distribution becomes the two-peak distribution. We remark, as the actual experimental results shown in Ref. [20] that the intensities of the distributions in (b) and (c) are smaller ($\sim 10^{-5}$) than the one in (a).

6.3 Testing the birefringence of the crystal

In this section, we see the hypothesis testing by the WVA through the experiment by Ritchie et al. [20]. We consider the testing problem to distinguish whether the crystal is birefringent or not in the optical setup. We have mathematically described the birefringence by fixing the observable $\hat{A} = \lambda_H |H\rangle\langle H| + \lambda_V |V\rangle\langle V|$ with the eigenvalues such as $\lambda_H \neq \lambda_V$ which means the refraction factors are different between the horizontal and the vertical polarization. We can set the hypotheses as

the null hypothesis H_0: the crystal is not birefringent, i.e., $\lambda_H = \lambda_V$,
the alternative hypothesis H_1: the crystal is birefringent, i.e., $\lambda_H \neq \lambda_V$.

We note that, in both hypotheses, the interaction, i.e., the refraction itself is still present ($g \neq 0$), which is different point from the previous problem considered in Chap. 5. To compare the detection capabilities of the conventional measurement and the weak measurement, we calculate each testing power (4.3) from the distributions $f_c(x|\lambda_H, \lambda_V) = f_c(x)$ for the conventional measurement and $f_w(x|\lambda_H, \lambda_V) = f_w(x)$ for the weak measurement. In this experimental setup, the power is given as the probability to determine exactly when the crystal is indeed birefringent.

Above all, we consider the test function. In the previous chapter, we have employed

$$d(x) = \begin{cases} 0 & \text{if } |x|/\sigma < c, \\ r & \text{if } |x|/\sigma = c, \\ 1 & \text{if } |x|/\sigma > c. \end{cases} \tag{6.10}$$

as the test function for the testing problem to investigate whether the interaction is present ($g \neq 0$) or not ($g = 0$). However, we have to pay attention that this test function is not suitable for the current test problem as it is because the beam position x can be deviated from the range of the initial beam waist w_0 by the refraction, although the null hypothesis test is true. To apply this test function to the current testing problem, we

need to adjust the final probe distribution by shifting $g\lambda_+ := g(\lambda_H + \lambda_V)/2$. 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 obtain the value of $g\lambda_+$ by the preparatory experiment without postselection to just monitor the refraction by the crystal. Thus, we can adjust the final probe distribution by a translation such as $y \to y + g\lambda_+$ in practical cases.

This adjustment can be given as the unitary operator $\exp[ig\lambda_+\hat{I} \otimes \hat{p}]$, where $\hat{I}$ is the identity operator of the measured system. The total unitary operator combining the two unitary operators given by the Hamiltonian (2.3) and by the adjustment becomes

$$\hat{U}_t = \exp[ig\lambda_+\hat{I} \otimes \hat{p}] \exp[-ig\hat{A} \otimes \hat{p}] = \exp[-ig(\hat{A} - \lambda_+\hat{I}) \otimes \hat{p}] = \exp[-ig\hat{A}_t \otimes \hat{p}], \quad (6.11)$$

where $\hat{A}_t := \lambda_- (|H\rangle\langle H| - |V\rangle\langle V|)$ is the total observable and $\lambda_- := (\lambda_H - \lambda_V)/2$ is its eigenvalue. The total weak value becomes

$$\langle \hat{A}_t \rangle_w = \frac{\langle f|\hat{A}_t|i\rangle}{\langle f|i\rangle} = \lambda_- \frac{\cos(\theta_i + \theta_f)}{\cos(\theta_i - \theta_f)}. \quad (6.12)$$

Furthermore, we can rewrite our two hypotheses as

the null hypothesis H_0: the crystal is not birefringent, i.e., $\lambda_- = 0$,
the alternative hypothesis H_1: the interaction would be present, i.e., $\lambda_- \neq 0$.

To calculate the testing powers of each measurement, we use the adjusted distributions:

$$f_c^{\mathrm{adj}}(y|\lambda_-) = f_{\mathrm{nps}}(y + g\lambda_+|\lambda_H, \lambda_V) = \frac{1}{\sqrt{2\pi w_0^2}} \Big(\cos^2\theta_i \; e^{-\frac{(y-g\lambda_-)^2}{2\sigma^2}} + \sin^2\theta_i \; e^{-\frac{(y+g\lambda_-)^2}{2\sigma^2}} \Big), \quad (6.13)$$

for the conventional measurement and

$$\begin{aligned} & f_w^{\mathrm{adj}}(y|\lambda_-) \\ &= f_{\mathrm{ps}}(y + g\lambda_+|\lambda_H, \lambda_V) \\ &= \frac{\cos^2\theta_i \cos^2\theta_f \; e^{-\frac{(y-g\lambda_-)^2}{2\sigma^2}} + \sin^2\theta_i \sin^2\theta_f \; e^{-\frac{(y+g\lambda_-)^2}{2\sigma^2}} + \frac{1}{2}\sin 2\theta_i \sin 2\theta_f \; e^{-\frac{y^2+g^2\lambda_-^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}\Big(\cos^2\theta_i \cos^2\theta_f + \sin^2\theta_i \sin^2\theta_f + \frac{1}{2}\sin 2\theta_i \sin 2\theta_f \; e^{-\frac{g^2\lambda_-^2}{2\sigma^2}} \Big)}. \end{aligned} \quad (6.14)$$

for the weak measurement.

In the previous chapter, we have proved that the test function (6.10) gives the UMPU test for the previous testing problem of $H_0 : g = 0$ and $H_1 : g \neq 0$. Because the distribution functions (6.13) and (6.14) are functions of the product $g\lambda_-$, the test function (6.10) is also the UMPU test for the current testing problem $H_0 : \lambda_- = 0$ and $H_1 : \lambda_- \neq 0$.

The testing power of the conventional measurement is calculated as

$$\beta_c(\lambda_-) = 1 - \frac{1}{2}\left(\mathrm{erf}\left[\frac{c\sigma - g\lambda_-}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c\sigma + g\lambda_-}{\sqrt{2\sigma^2}}\right]\right), \tag{6.15}$$

and the one of the weak measurement is

$$\beta_w(\lambda_-) = 1 - \frac{\begin{array}{r}(\cos^2\theta_i\cos^2\theta_f + \sin^2\theta_i\sin^2\theta_f)\left(\mathrm{erf}\left[\frac{c\sigma-g\lambda_-}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c\sigma+g\lambda_-}{\sqrt{2\sigma^2}}\right]\right) \\ + \sin 2\theta_i \sin 2\theta_f \; e^{-\frac{g^2\lambda_-^2}{2\sigma^2}} \mathrm{erf}\left[\frac{c}{\sqrt{2}}\right]\end{array}}{2(\cos^2\theta_i\cos^2\theta_f + \sin^2\theta_i\sin^2\theta_f) + \sin 2\theta_i \sin 2\theta_f \; e^{-\frac{g^2\lambda_-^2}{2\sigma^2}}}. \tag{6.16}$$

From these powers, we can obtain the relation

$$\frac{1-\beta_w(\lambda_-)}{1-\beta_c(\lambda_-)} - 1 = \frac{\sin 2\theta_i \sin 2\theta_f \; e^{-\frac{g^2\lambda_-^2}{2\sigma^2}}\left(\frac{2\mathrm{erf}\left[c/\sqrt{2}\right]}{\mathrm{erf}\left[(c\sigma-g\lambda_-)/\sqrt{2\sigma^2}\right]+\mathrm{erf}\left[(c\sigma+g\lambda_-)/\sqrt{2\sigma^2}\right]} - 1\right)}{2(\cos^2\theta_i\cos^2\theta_f + \sin^2\theta_i\sin^2\theta_f) + \sin 2\theta_i \sin 2\theta_f \; e^{-\frac{g^2\lambda_-^2}{2\sigma^2}}}, \tag{6.17}$$

which implies the inequality

$$\beta_w(\lambda_-) \geq \beta_c(\lambda_-), \tag{6.18}$$

under the condition that the pair of θ_i and θ_f satisfies that

$$C(\theta_i, \theta_f) := \sin 2\theta_i \sin 2\theta_f \leq 0. \tag{6.19}$$

Here we have used the inequality

$$\frac{2\mathrm{erf}\left[c/\sqrt{2}\right]}{\mathrm{erf}\left[(c\sigma - g\lambda_-)/\sqrt{2\sigma^2}\right] + \mathrm{erf}\left[(c\sigma + g\lambda_-)/\sqrt{2\sigma^2}\right]} - 1 > 0, \tag{6.20}$$

which holds for $g\lambda_- \neq 0$ and has been shown in Sec. 5.3. We note that the condition (6.19) coincides with negativity of the third term in the numerator of Eq. (6.14) generated

Table 6.1: Condition $C(\theta_i, \theta_f)$ and the weak value are displayed with the assigned values of θ_i and θ_f used in the actual experiment [20]. In each case, $\theta_i = \pi/4$.

	θ_f	$C(\theta_i, \theta_f)$	$\left\|\langle \hat{A}_t \rangle_w / \lambda_- \right\|^2$
(a)	$\pi/4$	1	0
(b)	$3\pi/4 + 2.2 \times 10^{-2}$	-0.999	2065
(c)	$3\pi/4$	-1	indeterminate

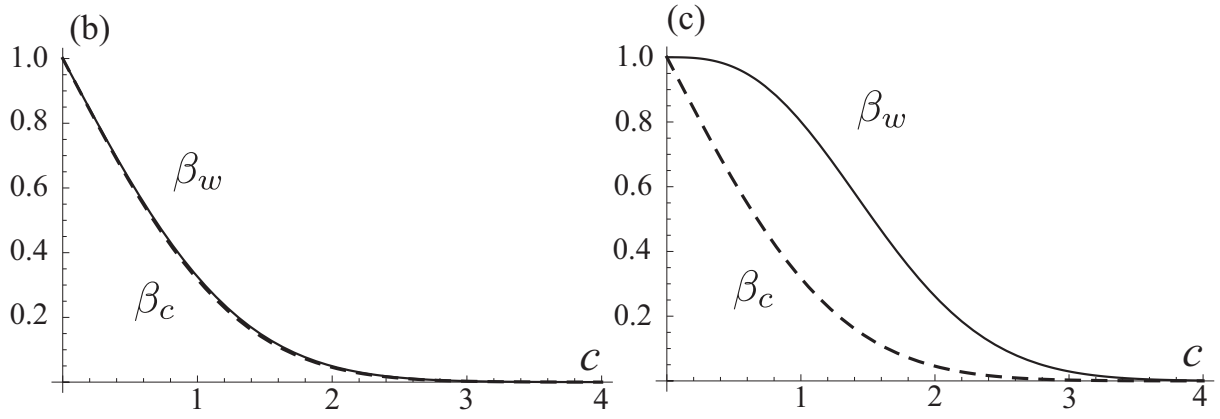


Figure 6.3: Powers β_w (solid line) and b_c (dashed line) plotted with the critical points c as the horizontal axis in the two cases: (b) $\theta_f = 3\pi/4 + 2.2 \times 10^{-2}$ and (c) $\theta_f = 3\pi/4$. The other parameters are fixed as $\theta_i = \pi/4$, $w_0 = 55\,\mu$m, and $g\lambda_- = 0.32, \mu$m.

by postselection which induce the amplification effect. We also remark that condition (6.19) is related to the inequality of a weak value (5.18) derived in the previous chapter as $|\langle \hat{A}_t \rangle_w|^2 \geq |\lambda_-|^2 \Leftrightarrow C(\theta_i, \theta_f) \leq 0$. Thus we have shown that the angles of the two polarizers in this experimental setup are the important factor that determines the case when the weak measurement is superior to the conventional measurement in terms of the testing power.

Table 6.1 represents the value of the condition $C(\theta_i, \theta_f)$ and the total weak value $|\langle \hat{A}_t \rangle_w / \lambda_-|^2$ for the angles of the two polarizer θ_i and θ_f which were used in the actual experiment [20]. In Fig. 6.2, we see two cases: (a) the roughly conventional measurement and (c) the weak measurement with the postselected state completely orthogonal to the preselected state. In case (a), the single Gaussian distribution is shown, so it is difficult to distinguish whether the crystal is birefringent or not. In contrast, the two-peak distribution appears in case (c), which enables us to clearly recognize that the crystal is birefringent. We compare (b) and (c), the both of which are the weak measurement

cases and the inequality (6.18) holds for the condition (6.19). Since the case (b) meets the AAV 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. [17]. On the other hand, the case (c) which does not satisfy the AAV approximation gives significant amplification in the point of view of the peak-to-peak distance [19].

Figure 6.3 shows the plots of powers given by the conventional measurement and the two weak measurement cases, (b) and (c). As shown in Fig. 6.3, the case (c) gives a remarkable effect to detect the birefringence of a crystal, whereas the plots of $b_{\rm ps}$ and $b_{\rm nps}$ are almost overlapped in the case (b).

6.4 Summary of this chapter

In the current task, we have studied the classic experiment [20], regarding it as testing the birefringence of the crystal. The experiment is a helpful example to consider the hypothesis testing with the WVA [60] because it is investigated outside the validity of approximation, especially the case of postselection completely orthogonal to preselection. We have shown that the weak measurement can be more powerful than the conventional measurement in the hypothesis testing to distinguish whether or not the crystal is birefringent for the specific experimental set-up. When the total weak value given by the angles of the two polarizers is larger than the eigenvalues of the total observable, the WVA has the advantage for the present problem. In particular, the pair of angles, which does not satisfy the AAV approximation $g|\langle\hat{A}\rangle_w| \ll 1$, gives the really powerful testing. According to the authors of Ref. [19], the amplification effect is rather striking when the approximation breaks down. Our conclusion obtained through the statistical analysis supports their view on the WVA.

In this chapter 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 chapter. In both cases, the testing function (6.10) gives the UMPU test and works well.

We note that we observe no data with a completely orthogonal pair of angles when the null hypothesis is really true. Practically, it is important to keep the two angles almost orthogonal but not quite, while the AAV approximation is not valid.

Chapter 7

Concluding remarks

7.1 Summary of this thesis

Throughout this thesis, we have studied the weak measurement and the weak-value amplification (WVA) proposed by Aharonov, Albert, and Vaidman (AAV) [17]. Especially the WVA is a promising technique for a precision measurement due to the amplification of the output of the probe[14–16]. In the original proposal, by assuming a weak coupling constant, they found that the weak value larger than eigenvalues of the measured system leads to the amplification. By this amplification effect, we suppose that the WVA will be usefull for extracting the information of the interaction between a measured system and a measuring probe in the indirect quantum measurement model. Some researchers have investigated a merit of the WVA in theoretical and experimental studies and have suggested the validity of the WVA. In this thesis, we have proposed and examined the two possibilities of merits for further developing the WVA.

First, we have derived the optimal probe wave function by the Lagrange multiplier method. In conventional discussion [32–34], the shift of the expectation value by the WVA with the Gaussian probe has an upper bound. We have shown that the optimal probe can maximize the amplified shift, which can be arbitrarily large as the weak value is, and gives the zero variance of final probe with a certain weak value, in principle.

Second, we have considered the WVA from the viewpoint of the statistical inference. Some researchers have already pointed out that the weak measurement is worse for the

parameter estimation that the conventional measurement due to the failure of the postselection. [40–44]. Especially, in this thesis, we have focused on the fact that the amplification effect dose not enlarge the Fisher information regardless of the success probability of the postselection. Therefore, we turn to another problem, statistical inference problem. Here we have considered the detection capability of the weak measurement through the problem that we determine whether or not the interaction is present. To address this question, we have used the hypothesis testing method [53–55]. The hypothesis testing is completely different from the parameter estimation as a task, although we are motivated by the dispute on the amplification effect in the parameter estimation problem. The standard strategy in the hypothesis testing is to make the probabilities of the type-1 and -2 errors small. The type-1 error represents the false alarm rate and the type-2 error is the false negative. In Chap. 5, we have established the appropriate test function for the current problem, which is classified as a UMPU test or a UMP test in a certain case. The test function is produced from the physical intuition that the interaction will be present if the output data is outside the initial probe fluctuation. Under the test function, without any approximation, we find that the weak measurement can be superior to the conventional measurement. It is shown that the merit of the WVA is the reduction of the possibility to miss the presence of the interaction more than the conventional measurement while keeping the probability of a misdetection, when the absolute value of the weak value is outside the range of the eigenvalues. Throughout discussion, we have considered the case that the number of the obtained data is infinitely large and we neglect the data loss by the failure of the postseleciton.

7.2 Discussion

The key property of the optimal probe wave function is the denominator of the function (3.1). Since the denominator cancel out the higher-order terms of the coupling constant in the unitary operator, which is understood as a back action of the measurement, the optimal probe wave function can give an arbitrarily large amplification. In practice, as we can see in Eq. (3.1), we have to know the coupling constant and the weak value before the experiment for realizing this optimal probe. To engineer the optimal probe, we use the pre-guessed value of the coupling constant and the calculated weak value from the experimental setup.

In the study of the weak measurement and the hypothesis testing, we have found that the WVA reduces the possibility to miss the presence of the interaction with a fixed false

alarm rate when the absolute value of the weak value is greater than the eigenvalues. We emphasize that our result provides a different physical intuition of the WVA from a conventional one that the weak value can be a figure of merit for the detection power. Eventually, the merit of the WVA is the increase of the detection power, which agrees with the previous intuition from the AAV [17]. Additionally we have shown that this conclusion remains true even if the measurement data suffers the additive white Gaussian noise. Also we have considered the situation that the null hypothesis has a small interval in Appendix D.2. There we have shown that when the type-2 error is more serious than the type-1 error, the weak measurement with a certain weak value has an advantage. In this discussion, we have supposed that the measured system is the two-state system and that the initial wave function of the measuring probe is Gaussian.

In Chap. 6, we have exemplified the optical experiment using a single birefringent crystal demonstrated by Ritchie *et al.* [20], which can be regraded as the testing problem for distinguishing whether or not the crystal has birefringence. Consequently, we have shown that the weak measurement is advantageous for this testing problem as well as the discussion in Chap. 5. Especially, we have found that when the AAV approximation ($g|\langle\hat{A}\rangle_w| \ll 1$) breaks down by the large weak value, the WVA can be more beneficial. This fact supports the view of the authors of Ref. [19].

7.3 Outlook

Recently, researchers dispute about a technical utility of the WVA. Some of them criticize the WVA for the demerit by the postselection. On the other hand, there are theoretical supports for the WVA and several experiments have shown the WVA.

Through the discussion of the optimal probe wave function, we find that the preparing initial probe other than Gaussian would be significant. To experimentally do it, the spatial phase and amplitude modulation are needed, which has been studied [65]. It might be possible to realize the optimal probe wave function in the future. Also some researchers have researched the weak measurement with the Laguerre-Gaussian [66–68] and found effectiveness. An alteration of initial probe is the one of the methods for developing the weak measurement. The optimal probe wave function is a one of the proposal for probe engineering in the WVA. In practice, it is one of the means to make appropriate probe wave function corresponding to the experimental environment and the purpose in reference to the optimal one.

In Chap. 5,we have found that the WVA is advantageous for the interaction detection by using the hypothesis testing. Throughout this thesis, we assume that the number of the obtained data is infinitely large. In a real experiment, the number of the obtainable data is finite. Then we should establish the formalism for the testing in the finite data with the weak measurement to more practical study. Meanwhile, there are many methods of the statistical inference other than the hypothesis testing. For instance, the Bayes factor has been studied as the alternative method of the hypothesis testing [69, 70]. In the hypothesis testing, we postulate that the tested parameter is completely unknown. In the Bayes factor, we use a prior probability distribution, which expresses our preliminary knowledge of the parameter that we want to test. In actual experiment, we have somewhat knowledge of the parameter on ahead. The weak measurement with the Bayes factor will be give a more powerful tool for a precise testing in a experiment.

While we have shown its validity thorough the optical experiment with a single birefringent crystal in Chap. 6, this result suggests a possibility that the proposed hypothesis testing method with the weak measurement can apply for other experiments, for example, the observation of the spin Hall effect of light [21], the accurate sensing of the tilted mirror in the Sagnac interferometer [22]. Here we note that to apply the proposed hypothesis testing method, some assumptions are needed to be satisfied that the initial probe is of the Gaussian profile, the measured system is two-state system, and an experiment is for measuring probe position. We need to set up the proper testing method corresponding to each experiment. In addition, the discussion about the data loss problem due to the postselection remains, which is partially studied in Appendix D.3. Furthermore, we have to care a case that we cannot obtain any data, which leads to the problem of the null result. This problem remains an open problem. We know how to treat the null result in the projective measurement [71, 72]. As for the null result in the weak measurement, the measurement protocol named the null weak value is proposed in Refs [73–75] for improved the signal-noise ratio, which will be of some help to address the problem.

The author hopes that the two theoretical proposals, the optimal probe wave function and the hypothesis method with the weak measurement, will contribute not only to develop the WVA and the precise measurement but also to discover a new physical phenomenon.

Appendix A

Derivation of amplified shift depending on initial probe

Here we derive Eq. (2.35), which shows that the expectation value depends on the initial prove wave function. For simplicity, we assume that the initial probe wave function $\tilde{\psi}(p)$ is an even pure real function and converges to 0 at $x \to \pm\infty$. For the calculation, we use the calculation formula as

$$B^*(p)B'(p) = -ig\mathrm{Re}A_w + \frac{1}{2}\left[|B(p)|^2\right]', \tag{A.1}$$

where the prime represents the differentiation of p, which is given by

$$\begin{aligned}
&\left[|B(p)|^2\right]' = B^{*\prime}(p)B(p) + B^*(p)B'(p) \\
=&(-g\sin gp + ig\langle\hat{A}\rangle_w^*\cos gp)(\cos gp - i\langle\hat{A}\rangle_w\sin gp) + B^*(p)B'(p) \\
=&-g\sin gp\cos gp + ig\langle\hat{A}\rangle_w^*(1-\sin^2 gp) \\
&+ ig\langle\hat{A}\rangle_w(1-\cos^2 gp) + g|\langle\hat{A}\rangle_w|^2\sin gp\cos gp + B^*(p)B'(p) \\
=&2ig\mathrm{Re}\langle\hat{A}\rangle_w + (\cos gp + i\langle\hat{A}\rangle_w^*\sin gp)(-g\sin gp - ig\langle\hat{A}\rangle_w\cos gp) + B^*(p)B'(p) \\
=&2ig\mathrm{Re}\langle\hat{A}\rangle_w + 2B^*(p)B'(p).
\end{aligned} \tag{A.2}$$

We see the expectation value of the final probe position which is

$$\langle\hat{x}\rangle_w = \frac{\langle\psi_w|\hat{x}|\psi_w\rangle}{\langle\psi_w|\psi_w\rangle} = \frac{\int dp\langle\psi_w|p\rangle\left(i\frac{\partial}{\partial p}\right)\langle p|\psi_w\rangle}{\int dp\langle\psi_w|p\rangle\langle p|\psi_w\rangle} = \frac{i\int dp\left[B(p)\tilde{\psi}(p)\right]^*\left[B(p)\tilde{\psi}(p)\right]'}{\int dp|B(p)\tilde{\psi}(p)|^2}. \tag{A.3}$$

The denominator is

$$\int dp|B(p)\tilde{\psi}(p)|^2 = \frac{1}{2}(1+|\langle\hat{A}\rangle_w|^2) + \frac{1}{2}(1-|\langle\hat{A}\rangle_w|^2)\int dp\cos 2gp|\tilde{\psi}(p)|^2 \tag{A.4}$$

The numerator of this equation becomes

$$\begin{aligned}
&i\int dp\left[B(p)\tilde{\psi}(p)\right]^*\left[B(p)\tilde{\psi}(p)\right]' = i\int dp\left[B^*(p)B'(p)|\tilde{\psi}(p)|^2 + |B(p)|^2\psi^*(p)\psi'(p)\right]\\
=&i\int dp\left[-ig\mathrm{Re}\langle\hat{A}\rangle_w + \frac{1}{2}\left(|B(p)|^2\right)'\right]|\tilde{\psi}(p)|^2 + i\int dp|B(p)|^2\tilde{\psi}^*(p)\tilde{\psi}'(p)\\
=&g\mathrm{Re}\langle\hat{A}\rangle_w\int dp|\tilde{\psi}(p)|^2 + \frac{i}{2}\int dp\left(|B(p)|^2\right)'|\tilde{\psi}(p)|^2 + i\int dp|B(p)|^2\tilde{\psi}^*(p)\tilde{\psi}'(p)\\
=&g\mathrm{Re}\langle\hat{A}\rangle_w + \frac{i}{2}\int dp\left[\left(|B(p)|^2|\tilde{\psi}(p)|^2\right)' - |B(p)|^2\left(|\tilde{\psi}(p)|^2\right)'\right] + i\int dp|B(p)|^2\tilde{\psi}^*(p)\tilde{\psi}'(p)\\
=&g\mathrm{Re}\langle\hat{A}\rangle_w + \int dp|B(p)|^2\left[i\frac{\tilde{\psi}^*(p)\tilde{\psi}'(p) - \tilde{\psi}^{*\prime}(p)\tilde{\psi}(p)}{2}\right] + \frac{i}{2}\int dp\left[|B(p)\tilde{\psi}(p)|^2\right]'\\
=&g\mathrm{Re}\langle\hat{A}\rangle_w + \mathrm{Re}\left[i\int dp|B(p)|^2\tilde{\psi}^*(p)\tilde{\psi}'(p)\right] + \frac{i}{2}\int dp\left[|B(p)\tilde{\psi}(p)|^2\right]'.
\end{aligned} \tag{A.5}$$

Because the initial probe wave function $\tilde{\psi}(p)$ is a pure real function and converges to 0 at $x \to \pm\infty$, the numerator is $g\mathrm{Re}\langle\hat{A}\rangle_w$. Thus the expectation value of the final probe position is given as

$$\langle\hat{x}\rangle_w = g\frac{2\mathrm{Re}\langle\hat{A}\rangle_w}{(1+|\langle\hat{A}\rangle_w|^2) + (1-|\langle\hat{A}\rangle_w|^2)\int dp\cos 2gp|\tilde{\psi}(p)|^2}. \tag{A.6}$$

Appendix B

Cramér-Rao Inequality and Fisher Information

B.1 Derivation of Cramér-Rao inequality

Here we show the Cramér-Rao inequality

$$\langle(\tilde{\theta}(x) - \theta)^2\rangle \geq 1/I_n(\theta), \tag{B.1}$$

where $I(\theta)$ is the classical Fisher information defined as

$$I_n(\theta) = \int dx\, [\partial_\theta \log f(x|\theta)]^2 f(x|\theta) \tag{B.2}$$

and n means the number of the data. In the following, we use the notations $l(x|\theta) := \log f(x|\theta)$ and a prime as a differentiation by the θ.

We assume that the integration and differentiation are commutable, which means

$$\partial_\theta \int dx\, f(x|\theta) = \int dx\, \partial_\theta f(x|\theta) = 0. \tag{B.3}$$

From this equation, we can see

$$\int dx\, \theta l'(x|\theta) f(x|\theta) = \theta \int dx\, \partial_\theta f(x|\theta) = 0. \tag{B.4}$$

Here we use an unbiased estimator $\tilde{\theta}(x)$ that

$$\theta = \langle \tilde{\theta}(x) \rangle = \int dx\, \tilde{\theta}(x) f(x|\theta), \tag{B.5}$$

which gives

$$1 = \partial_\theta \int dx\, \tilde{\theta}(x) f(x|\theta) = \int dx\, \tilde{\theta}(x) l'(x|\theta) f(x|\theta). \tag{B.6}$$

From Eqs. (B.4) and (B.6), we obtain

$$1 = \int dx\, (\tilde{\theta}(x) - \theta) l'(x|\theta) f(x|\theta). \tag{B.7}$$

By using the Cauchy-Schwarz inequality, we can show the Cramér-Rao inequality (B.1) as

$$\begin{aligned} 1 &= \left(\int dx\, (\tilde{\theta}(x) - \theta) l'(x|\theta) f(x|\theta) \right)^2 \\ &\leq \left(\int dx\, (\tilde{\theta}(x) - \theta)^2 f(x|\theta) \right) \left(\int dx\, [l'(x|\theta)]^2 f(x|\theta) \right) \\ &= \langle (\tilde{\theta}(x) - \theta)^2 \rangle I_n(\theta). \end{aligned} \tag{B.8}$$

Here we see $l''(x|\theta)$ as

$$l''(x|\theta) = \partial_\theta \left(\frac{\partial_\theta f(x|\theta)}{f(x|\theta)} \right) = \frac{\partial_\theta^2 f(x|\theta)}{f(x|\theta)} - [l'(x|\theta)]^2. \tag{B.9}$$

Again, we use the assumption that the integration and differentiation are commutable, which gives

$$\partial_\theta^2 \int dx\, f(x|\theta) = \int dx\, \partial_\theta^2 f(x|\theta) = 0 \tag{B.10}$$

for Eq. (B.3). From Eqs. (B.9) and (B.10), we get another form of the Fisher information as

$$\begin{aligned} \int dx\, l''(x|\theta) f(x|\theta) &= \int dx\, \frac{\partial_\theta^2 f(x|\theta)}{f(x|\theta)} f(x|\theta) - \int dx\, [l'(x|\theta)]^2 f(x|\theta) \\ \Rightarrow \; I_n(\theta) &= - \int dx\, l''(x|\theta) f(x|\theta). \end{aligned} \tag{B.11}$$

Next we show $I_n(\theta) = nI_1(\theta)$ when the n samples $x_1, \cdots, x_n$ comes from the independent identical distributions. We can see the joint probability distribution function

$$f_n(x_1, \cdots, x_n|\theta) = \prod_{i=1}^{n} f_1(x_i|\theta), \tag{B.12}$$

which implies

$$l_n''(x_1, \cdots, x_n|\theta) = \sum_{i=1}^{n} l_1''(x_1|\theta). \tag{B.13}$$

Because of Eq. (B.11), the Fisher information $I_n(\theta)$ for the joint probability distribution function can be denoted as

$$\begin{aligned} I_n(\theta) &= -\int \cdots \int dx_1 \cdots dx_n \, l_n''(x_1, \cdots, x_n|\theta) f_n(x_1, \cdots, x_n|\theta) \\ &= -\int \cdots \int dx_1 \cdots dx_n \sum_{i=1}^{n} l_1''(x_1|\theta) \prod_{i=1}^{\infty} f_1(x_i|\theta) \\ &= -\sum_{i=1}^{n} \int dx_1 \, l_1''(x_1|\theta) f_1(x_i|\theta) \\ &= nI_1(\theta). \end{aligned} \tag{B.14}$$

Combining Eq. (B.1) with Eq. (B.14), we can probe Eq. (2.43).

B.2 Calculation of Fisher information

B.2.1 Probe in position space

First we calculate the Fisher information for the conventional measurement providing the final probe distribution (2.45):

$$f^c(x|g) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-g)^2}{2\sigma^2}}. \tag{B.15}$$

Because the logarithmic of $f^c(x|g)$ is

$$\log f^c(x|g) = -\frac{1}{2} \log 2\pi\sigma^2 - \frac{(x-g)^2}{2\sigma^2}, \tag{B.16}$$

the derivative of which becomes

$$\partial_g \log f^c(x|g) = \frac{x-g}{\sigma^2}. \tag{B.17}$$

Thus the Fisher information for the conventional measurement is

$$I^c(g) = \frac{1}{\sqrt{2\pi\sigma^2}} \int dx \left(\frac{x-g}{\sigma^2}\right)^2 e^{-\frac{(x-g)^2}{2\sigma^2}} = \frac{1}{\sigma^2}. \tag{B.18}$$

Next we consider the Fisher information for the weak measurement giving the final probe distribution (2.47):

$$f^w_\epsilon(x|g) = \frac{1}{2\sqrt{2\pi\sigma^2}} \frac{e^{-\frac{(x-g)^2}{2\sigma^2}} + e^{-\frac{(x+g)^2}{2\sigma^2}} + 2\epsilon e^{-\frac{x^2+g^2}{2\sigma^2}}}{1+\epsilon e^{-\frac{g^2}{2\sigma^2}}}, \tag{B.19}$$

where $\epsilon = \pm 1$. The logarithmic derivative of $f^w_\epsilon(x|g)$ is calculated as

$$\begin{aligned}
\partial_g \log f^w_\epsilon(x|g) &= \frac{\frac{x-g}{\sigma^2} e^{-\frac{(x-g)^2}{2\sigma^2}} - \frac{x+g}{\sigma^2} e^{-\frac{(x+g)^2}{2\sigma^2}} + 2\epsilon(-\frac{g}{\sigma^2}) e^{-\frac{x^2+g^2}{2\sigma^2}}}{e^{-\frac{(x-g)^2}{2\sigma^2}} + e^{-\frac{(x+g)^2}{2\sigma^2}} + 2\epsilon e^{-\frac{x^2+g^2}{2\sigma^2}}} + \frac{\epsilon \frac{g}{\sigma^2} e^{-\frac{g^2}{2\sigma^2}}}{1+\epsilon e^{-\frac{g^2}{2\sigma^2}}} \\
&= \frac{1}{\sigma^2}\left[\frac{x(e^{-\frac{(x-g)^2}{2\sigma^2}} - e^{-\frac{(x+g)^2}{2\sigma^2}})}{e^{-\frac{(x-g)^2}{2\sigma^2}} + e^{-\frac{(x+g)^2}{2\sigma^2}} + 2\epsilon e^{-\frac{x^2+g^2}{2\sigma^2}}} - g + g\frac{\epsilon e^{-\frac{g^2}{2\sigma^2}}}{1+\epsilon e^{-\frac{g^2}{2\sigma^2}}}\right] \\
&= \frac{1}{g\sigma^2}\left[x\frac{gh(x|g)}{f^w_\epsilon(x|g)} - a\right],
\end{aligned} \tag{B.20}$$

where

$$h(x|g) := \frac{1}{2\sqrt{2\pi\sigma^2}} \frac{e^{-\frac{(x-g)^2}{2\sigma^2}} - e^{-\frac{(x+g)^2}{2\sigma^2}}}{1+\epsilon e^{-\frac{g^2}{2\sigma^2}}}, \tag{B.21}$$

$$a := \frac{g^2}{1+\epsilon e^{-\frac{g^2}{2\sigma^2}}}. \tag{B.22}$$

The squared one is

$$[\partial_g \log f^w_\epsilon(x|g)]^2 = \frac{1}{g^2\sigma^4}\left[x^2\frac{g^2h^2(x|g)}{(f^w_\epsilon(x|g))^2} - 2ax\frac{gh(x|g)}{f^w_\epsilon(x|g)} + a^2\right]. \tag{B.23}$$

Therefore we obtain the equation

$$g^2\sigma^4 I^w_\epsilon(g) = g^2\int dx\, x^2 \frac{h^2(x|g)}{f^w_\epsilon(x|g)} - 2ag\int dx\, xh(x|g) + a^2. \tag{B.24}$$

The integration of the second term is

$$\int dx\, xh(x|g) = \frac{a}{2g^2\sqrt{2\pi\sigma^2}} \int dx\, x(e^{-\frac{(x-g)^2}{2\sigma^2}} - e^{-\frac{(x+g)^2}{2\sigma^2}}) = \frac{a}{2g^2}[g-(-g)] = \frac{a}{g}. \quad \text{(B.25)}$$

Because of $\epsilon = \pm 1$, the integration of the first term is calculated as

$$\begin{aligned}
\int dx\, x^2 \frac{h^2(x|g)}{f_\epsilon^w(x|g)} &= \frac{a}{2g^2\sqrt{2\pi\sigma^2}} \int dx\, x^2 \frac{(e^{-\frac{(x-g)^2}{2\sigma^2}} - e^{-\frac{(x+g)^2}{2\sigma^2}})^2}{e^{-\frac{(x-g)^2}{2\sigma^2}} + e^{-\frac{(x+g)^2}{2\sigma^2}} + 2\epsilon e^{-\frac{x^2+g^2}{2\sigma^2}}} \\
&= \frac{a}{2g^2\sqrt{2\pi\sigma^2}} \int dx\, x^2 e^{-\frac{x^2+g^2}{2\sigma^2}} \frac{(e^{\frac{xg}{\sigma^2}} - e^{-\frac{xg}{\sigma^2}})^2}{e^{\frac{xg}{\sigma^2}} + e^{-\frac{xg}{\sigma^2}} + 2\epsilon} \\
&= \frac{a}{2g^2\sqrt{2\pi\sigma^2}} \int dx\, x^2 e^{-\frac{x^2+g^2}{2\sigma^2}} \frac{(e^{\frac{xg}{2\sigma^2}} - e^{-\frac{xg}{2\sigma^2}})^2 (e^{\frac{xg}{2\sigma^2}} + e^{-\frac{xg}{2\sigma^2}})^2}{(e^{\frac{xg}{2\sigma^2}} + \epsilon e^{-\frac{xg}{2\sigma^2}})^2} \\
&= \frac{a}{2g^2\sqrt{2\pi\sigma^2}} \int dx\, x^2 e^{-\frac{x^2+g^2}{2\sigma^2}} (e^{\frac{xg}{2\sigma^2}} - \epsilon e^{-\frac{xg}{2\sigma^2}})^2 \\
&= \frac{a}{2g^2\sqrt{2\pi\sigma^2}} \int dx\, x^2 (e^{-\frac{(x-g)^2}{2\sigma^2}} + e^{-\frac{(x+g)^2}{2\sigma^2}} - 2\epsilon e^{-\frac{x^2+g^2}{2\sigma^2}}) \\
&= \frac{a}{2g^2\sqrt{2\pi\sigma^2}} \int dx\, [(x+g)^2 + (x-g^2) - 2\epsilon e^{-\frac{g^2}{2\sigma^2}} x^2] e^{-\frac{x^2}{2\sigma^2}} \\
&= \frac{a}{g^2\sqrt{2\pi\sigma^2}} \int dx\, [(1-\epsilon e^{-\frac{g^2}{2\sigma^2}})x^2 + g^2] e^{-\frac{x^2}{2\sigma^2}} \\
&= \frac{a(1-\epsilon e^{-\frac{g^2}{2\sigma^2}})}{g^2} \frac{1}{\sqrt{2\pi\sigma^2}} \int dx\, x^2 e^{-\frac{x^2}{2\sigma^2}} + a \\
&= \frac{a(1-\epsilon e^{-\frac{g^2}{2\sigma^2}})}{g^2}\sigma^2 + a. \qquad \text{(B.26)}
\end{aligned}$$

Hence the Fisher information for the weak measurement becomes

$$I_\epsilon^w(g) = \frac{1}{g^2\sigma^4}\left[g^2\left(\frac{a(1-\epsilon e^{-\frac{g^2}{2\sigma^2}})}{g^2}\sigma^2 + a\right) - 2ag\frac{a}{g} + a^2\right] = \frac{1}{\sigma^2}\frac{1+\epsilon\frac{g^2}{\sigma^2}e^{-\frac{g^2}{2\sigma^2}} - e^{-\frac{g^2}{\sigma^2}}}{(1+\epsilon e^{-\frac{g^2}{2\sigma^2}})^2}. \quad \text{(B.27)}$$

B.2.2 Probe in momentum space

When the interaction Hamiltonian is given as $\hat{H}_{\rm int} = g\delta(t)\hat{A}\otimes\hat{p}$, the probe distribution in the momentum space does not change by interaction without postselection. We cannot obtain the information of the coupling constant from the final probe distribution in the momentum space given by the conventional measurement. Therefore we consider only the case of the weak measurement. Here we fix $\theta_i = \theta_f = \frac{\pi}{2}$ and $\phi = \phi_f - \phi_i$ for analytically

calculation. The probe distribution in the momentum space after the postselection (2.21) becomes

$$f(p|g) = \frac{1+\cos(2gp-\phi)}{1+\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}} \left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{2}} e^{-2\sigma^2 p^2}, \tag{B.28}$$

and the success probability of the postselection (2.41) is

$$P = \frac{1+\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}}{2}. \tag{B.29}$$

The logarithmic derivative of $f(p|g)$ is calculated as

$$\begin{aligned}
\partial_g \log f(p|g) &= \frac{-2p\sin(2gp-\phi)}{1+\cos(2gp-\phi)} - \frac{-\frac{g}{\sigma^2}\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}}{1+\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}} \\
&= \frac{\sqrt{2}}{\sigma}\left[\frac{-\sqrt{2}\sigma p\sin(2gp-\phi)}{1+\cos(2gp-\phi)} + \frac{\frac{g}{\sqrt{2}\sigma}\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}}{1+\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}}\right] \\
&= \frac{\sqrt{2}}{\sigma}\left[\frac{-\sqrt{2}\sigma p\sin(2gp-\phi)}{1+\cos(2gp-\phi)} + G\right],
\end{aligned} \tag{B.30}$$

where

$$G := \frac{\frac{g}{\sqrt{2}\sigma}\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}}{1+\cos\phi\, e^{-\frac{g^2}{2\sigma^2}}}. \tag{B.31}$$

The Fisher information multiplied by the success probability of the postselection is

$$\begin{aligned}
PI &= \int dp \frac{2}{\sigma^2}\left[\frac{-\sqrt{2}\sigma p\sin(2gp-\phi)}{1+\cos(2gp-\phi)} + G\right]^2 \frac{1+\cos(2gp-\phi)}{2}\left(\frac{2\sigma^2}{\pi}\right)^{\frac{1}{2}} e^{-2\sigma^2 p^2} \\
&= \int \frac{d\bar{p}}{\sqrt{2}\sigma}\frac{2}{\sigma^2}\left[\frac{-\bar{p}\sin(2\bar{g}\bar{p}-\phi)}{1+\cos(2\bar{g}\bar{p}-\phi)} + G\right]^2 \frac{1+\cos(2\bar{g}\bar{p}-\phi)}{2}\frac{\sqrt{2}\sigma}{\sqrt{\pi}} e^{-\bar{p}^2} \\
&= \frac{1}{\sqrt{\pi}\sigma^2}\int d\bar{p}\left[\frac{\bar{p}^2\sin^2(2\bar{g}\bar{p}-\phi)}{1+\cos(2\bar{g}\bar{p}-\phi)} - 2G\bar{p}\sin(2\bar{g}\bar{p}-\phi) + G^2\left(1+\cos(2\bar{g}\bar{p}-\phi)\right)\right] e^{-\bar{p}^2}
\end{aligned} \tag{B.32}$$

Here we have defined $\bar{p} = \sqrt{2}\sigma p,\ \bar{g} = g/\sqrt{2}\sigma$. The first term of the integrand becomes

$$\frac{\sin^2(2\bar{g}\bar{p}-\phi)}{1+\cos(2\bar{g}\bar{p}-\phi)} = \frac{1-\cos^2(2\bar{g}\bar{p}-\phi)}{1+\cos(2\bar{g}\bar{p}-\phi)} = 1-\cos(2\bar{g}\bar{p}-\phi). \tag{B.33}$$

Thus Eq. (B.32) is arranged as

$$\begin{aligned}
\sqrt{\pi}\sigma^2 PI &= \int d\bar{p}\,\bar{p}^2\,(1-\cos(2\bar{g}\bar{p}-\phi))\,e^{-\bar{p}^2} \\
&\quad - 2G\int d\bar{p}\,\bar{p}\sin(2\bar{g}\bar{p}-\phi)e^{-\bar{p}^2} + G^2\int d\bar{p}\,(1+\cos(2\bar{g}\bar{p}-\phi))\,e^{-\bar{p}^2} \\
&= \int d\bar{p}\,\bar{p}^2 e^{-\bar{p}^2} - \cos\phi\int d\bar{p}\,\bar{p}^2\cos(2\bar{g}\bar{p})e^{-\bar{p}^2} - 2G\cos\phi\int d\bar{p}\,\bar{p}\sin(2\bar{g}\bar{p})e^{-\bar{p}^2} \\
&\quad + G^2\left[\int d\bar{p}\,e^{-\bar{p}^2} + \cos\phi\int d\bar{p}\cos(2\bar{g}\bar{p})e^{-\bar{p}^2}\right] \\
&= \int d\bar{p}\,\bar{p}^2 e^{-\bar{p}^2} - \cos\phi\,e^{-\bar{g}^2}\int d\bar{p}\,\bar{p}^2 e^{-(\bar{p}-i\bar{g})^2} - 2G\cos\phi\,e^{-\bar{g}^2}\mathrm{Im}\int d\bar{p}\,\bar{p}e^{-(\bar{p}-i\bar{g})^2} \\
&\quad + G^2\left[\int d\bar{p}\,e^{-\bar{p}^2} + \cos\phi\,e^{-\bar{g}^2}\int d\bar{p}\,e^{-(\bar{p}-i\bar{g})^2}\right].
\end{aligned} \tag{B.34}$$

Since

$$\int d\bar{p}\,\bar{p}^2 e^{-(\bar{p}-i\bar{g})^2} = \sqrt{\pi}\left(\frac{1}{2}-\bar{g}^2\right),\ \int d\bar{p}\,\bar{p}e^{-(\bar{p}-i\bar{g})^2} = i\bar{g}\sqrt{\pi},\ \int d\bar{p}\,e^{-(\bar{p}-i\bar{g})^2} = \sqrt{\pi}, \tag{B.35}$$

we can calculate the integration as

$$\begin{aligned}
\sigma^2 PI &= \frac{1}{2} - \cos\phi\,e^{-\bar{g}^2}\left(\frac{1}{2}-\bar{g}^2\right) - 2G\cos\phi\,e^{-\bar{g}^2}\bar{g} + G^2\left(1+\cos\phi\,e^{-\bar{g}^2}\right) \\
&= \frac{1}{2} - \frac{1}{2}\cos\phi\,e^{-\bar{g}^2} + \bar{g}^2\cos\phi\,e^{-\bar{g}^2} - 2\frac{\bar{g}^2\cos^2\phi\,e^{-2\bar{g}^2}}{1+\cos\phi\,e^{-\bar{g}^2}} + \frac{\bar{g}^2\cos^2\phi\,e^{-2\bar{g}^2}}{1+\cos\phi\,e^{-\bar{g}^2}} \\
&= \frac{1}{2}(1-\cos\phi\,e^{-\bar{g}^2}) + \frac{\bar{g}^2\cos\phi\,e^{-\bar{g}^2} + \bar{g}^2\cos^2\phi\,e^{-2\bar{g}^2}}{1+\cos\phi\,e^{-\bar{g}^2}} - \frac{\bar{g}^2\cos^2\phi\,e^{-2\bar{g}^2}}{1+\cos\phi\,e^{-\bar{g}^2}} \\
&= \frac{1-\cos^2\phi\,e^{-2\bar{g}^2} + 2\bar{g}^2\cos\phi\,e^{-\bar{g}^2}}{2(1+\cos\phi\,e^{-\bar{g}^2})}.
\end{aligned} \tag{B.36}$$

Fig. B.1 plots PI in three cases, $\phi = 0$ for the postselection parallel to the preselection, $\phi = 3$ for the postslection almost orthogonal to the preselection, and $\phi = \pi$ for the postslection completely orthogonal to the preselection. As we can see from the figure, to measure the small coupling constant $\bar{g}$, we should tune the postslection almost orthogonal to the preselection, not completely. This result includes the failure of the postselection.

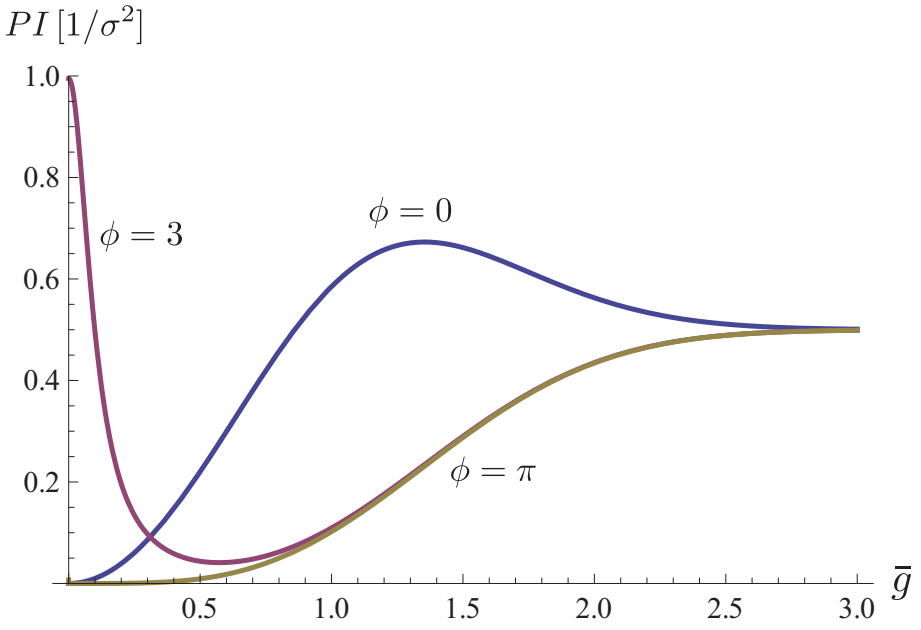


Figure B.1: Plots for the Fisher information multiplied by the success probability of the postselection PI in the three cases, the blue curve for $\phi = 0$, the red curve for $\phi = 3$, and the yellow curve for $\phi = \pi$. The vertical axis is normalized by $1/\sigma^2$, and the horizontal axis represents $\bar{g}$.

Appendix C

Shot Noise in Optical Experiment

C.1 Preface

In the main part of this thesis, we focus on the utility of the WVA in the aspect of the theoretical statistical inference. However, in practical experiment, various noises disturb the measurement. For example, in optical phase measurement, there is a photon shot noise derived from a photon number fluctuation. Here, we see that the shot noise is always larger than the variance of the probe distribution. Because the weak measurements are often demonstrated in optical experiment, we should take account of the shot noise when we consider the measurement accuracy. This topic was discussed in Ref. [36] which also showed a measurement limit of interaction in the WVA with the shot noise.

C.2 Example of experimental setup

Here we exemplify the optical phase measurement in the Mach-Zehnder interferometer as shown in Fig. C.1. We attempt to measure the frequency shift of the probe caused by slightly moved mirrors on corners, which is treated in momentum space as in the main part. The injected laser is separated into two optical paths, upper one and lower one, by the 50:50 beam splitter. We regard the which-path state as the measured system, then we note $|\uparrow\rangle$ and $|\downarrow\rangle$ as the states that the photon goes the upper and lower paths, respectively. We set a phase shifter on upper path, which induces the phase difference between the two states. We can fix the preselected state as $|i\rangle = (e^{i\phi/2}|\uparrow\rangle + e^{-i\phi/2}|\downarrow\rangle)/\sqrt{2}$, where the phase shift ϕ is symmetrized for convenience. By tuning the phase shiftier, we can

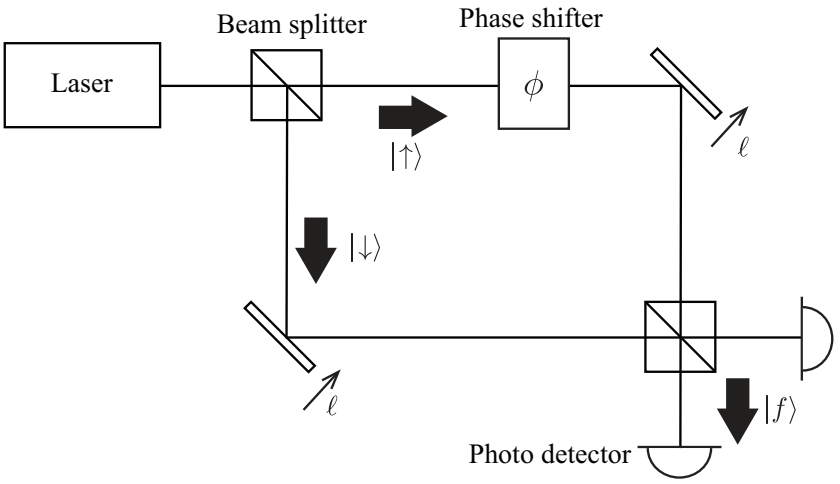


Figure C.1: Mach-Zehnder interferometer with slightly displaced mirrors which give the interaction

control the orthogonality between the pre- and postselected states. In both paths, we set slightly displaced mirrors which gives the interaction between the paths and the phase of beams. ℓ represents a small displacement of the mirrors . We can see that the interaction Hamiltonian is given as $\hat{H} = g\delta(t)\hat{A} \otimes \hat{\omega}$, where the interaction strength $g = \sqrt{2}\ell$, the observable $\hat{A} = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|$, and $\hat{\omega}$ is a frequency operator. After the interaction and second beam splitter which makes the postselected state as $|f\rangle = (|\uparrow\rangle - |\downarrow\rangle)/\sqrt{2}$, we observe the probe at a photo detector on an asymmetric output port. The port becomes almost dark by tuning the phase $\phi \approx 0$. The weak value in this optical setup is $\langle\hat{A}\rangle_w = -i\cot\phi/2$ and the frequency shift is given by the expectation value of the final probe $\langle\omega\rangle_w$ calculated as well as Eq. (2.24). From the shift, we can obtain the displacement of the mirrors ℓ. However, the final probe distribution has a variance $\langle\omega^2\rangle_w - \langle\omega\rangle_w^2$, which can be regraded as the frequency noise. Also, in optical experiment, the photon shot noise appears, which is the photon number fluctuation at the output port. Next, we evaluate the photon shot noise.

C.3 Evaluating shot noise

Here we use the notations $|\Psi\rangle$ as the final normalized probe state and $n(\omega)$ as a number of the photon at the output port for photon frequency ω. We assume that the output state is coherent in each frequency mode. From these, the frequency shift for a single

photon can be described as

$$\langle\omega\rangle_w = \int d\omega\, \omega |\langle\omega|\Psi_w\rangle|^2. \tag{C.1}$$

When we have N photons at output port in total, the averaged photon-number distribution for the frequency is given as

$$\overline{n(\omega)} = N|\langle\omega|\Psi_w\rangle|^2. \tag{C.2}$$

The expectation value of the frequency shift is

$$\langle\omega\rangle_w = \frac{1}{N}\int d\omega\, \omega\overline{n(\omega)}. \tag{C.3}$$

Here we take the fluctuation of the photon number noted as $\Delta n(\omega)$ into account, because the photon number $n(\omega)$ observed in a experiment fluctuates around its average $\overline{n(\omega)}$, i.e., $n(\omega) = \overline{n(\omega)} + \Delta n(\omega)$. The measured frequency shift becomes

$$\omega_{\text{measured}} = \frac{1}{N}\int d\omega\, \omega n(\omega) = \langle\omega\rangle_w + \Delta\omega, \tag{C.4}$$

where

$$\Delta\omega = \frac{1}{N}\int d\omega\, \omega\Delta n(\omega) \tag{C.5}$$

is the shot noise derived from the fluctuation of the photon number. For the coherence of the output probe, the photon number fluctuation $\Delta n(\omega)$ conforms the Poisson distribution. Thus the expectation value of the frequency shift fluctuation is $\langle\Delta\omega\rangle_p = 0$, where the subscript p represents the averaging over the Poisson distribution. The variance is

$$\begin{aligned}\text{Var}[\Delta\omega] = \langle(\Delta\omega)^2\rangle_p &= \frac{1}{N^2}\int d\omega \int d\omega' \omega\omega' \langle n(\omega)n(\omega')\rangle_p \\ &= \frac{1}{N^2}\int d\omega\, \omega^2\, \overline{n(\omega)} = \frac{1}{N}\langle\omega^2\rangle_w.\end{aligned} \tag{C.6}$$

Here we use the equation

$$\langle n(\omega)n(\omega')\rangle_p = \overline{n(\omega)}\delta(\omega - \omega'). \tag{C.7}$$

Hence the signal-to-noise ratio is given as

$$\mathrm{SNR} = \frac{|\langle\omega\rangle_w|}{\sqrt{\Delta\omega}} = \sqrt{N}\frac{|\langle\omega\rangle_w|}{\sqrt{\langle\omega^2\rangle_w}}, \tag{C.8}$$

which means that the shot-noise is always larger than the frequency noise.

This discussion can be applied for other optical setup, e.g. the Sagnac interferometer and the Michelson interferometer. Consequently, when we discuss the improvement of the signal-to-noise ratio or measurement accuracy of the weak measurement in an optical experiment, we need to consider the fluctuation of the photon number rather than the variance of the final probe distribution.

Appendix D

Further Study of Testing Method with Weak-Value Amplification

D.1 Preface

In Chap 6, we see the basic situation of the testing method with the weak-value amplification for the interaction detection problem. Here, we further discuss some extended situation. In Sec. D.2, we discuss the hypothesis testing in a small interval null hypothesis with the same physical setup considered in Chap. 5. This discussion provides more practical In Sec. D.3, we consider the proposed testing method with the data loss by postselection which is sometimes argued as the disadvantage of the weak-value amplification. Here we produce the makeshift model of the test function which include the risk of the data loss and analysis with the Lagrange multiplier method. This section is based on the Appendix C of Ref. [60].

D.2 Testing in small interval null hypothesis

In the previous chapter, we have set the null hypothesis (that the interaction is absent) as $g = 0$. However, there is following criticism for such a hypotheses set-up of the two-sided testing, i.e., $H_0 : g = 0$ and $H_1 : g \neq 0$. J. O. Berger (1985) said in his text book [56] that

> "Now it is unlikely that the null hypotheses is ever exactly true. Suppose, for instance, that $\theta = 10^{-10}$, which nonzero is probably meaning less difference

> form zero in most practical contexts." "What we are really interested in determining is whether or not the null hypothesis is approximately true (see Subsection 4.3.3). In Example 8, for instance, we might really be interested in detecting a difference of at least 10^{-3} from zero, in which case a better null hypothesis would be $H_0 : |\theta| \leq 10^{-3}$."

In response to this claim, here we suppose the following hypotheses;

> the null hypothesis H_0: the interaction would be absent, i.e., $|g| \leq \varepsilon$,
> the alternative hypothesis H_1: the interaction would be present, i.e., $|g| > \varepsilon$,

where ε is the real value depending on the experimental set-up, which will be small in practical case. Because the null hypothesis contains $g \neq 0$, the probabilities of the type-1 error are not same between the weak measurement and the conventional measurement. So, we need to produce a different way for comparison from the previous discussion. Here we propose the following cost function [1] as

$$r\Pr[\mathcal{E}_1] + (1-r)\Pr[\mathcal{E}_2], \tag{D.1}$$

where the real value $r(0 \leq r \leq 1)$ represents the weighting of the two error. In this section, we argue which measurement can more reduce the cost function under the testing function

$$d(x) = \begin{cases} 0 & \text{if } |x|/\sigma < c, \\ r & \text{if } |x|/\sigma = c, \\ 1 & \text{if } |x|/\sigma > c, \end{cases} \tag{D.2}$$

which is same as one used in Chap. 5.

We see the probabilities of the two errors in each measurement. Here, we introduce the interaction strengths g_1 and g_2 which satisfy $0 < |g_1| \leq \varepsilon < |g_2|$. The probability of the type-2 error in the weak measurement calculated in the previous chapter is still valid in the current testing problem. The probability of the type-2 error for the interaction g_2

[1] 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.

and the weak value $\langle \hat{A} \rangle_w$ is given by

$$\begin{aligned} &E_2(g_2, |\langle \hat{A} \rangle_w|^2) := \Pr[\mathcal{E}_2] \\ =&\frac{1}{2\mathcal{Z}} \left[(1 + |\langle \hat{A} \rangle_w|^2) \left(\mathrm{erf} \left[\frac{c\sigma - g_2}{\sqrt{2\sigma^2}} \right] + \mathrm{erf} \left[\frac{c\sigma + g_2}{\sqrt{2\sigma^2}} \right] \right) + 2(1 - |\langle \hat{A} \rangle_w|^2) e^{-\frac{g_2^2}{2\sigma^2}} \mathrm{erf} \left[\frac{c}{\sqrt{2}} \right] \right], \end{aligned} \tag{D.3}$$

which is derived by the integration of the distribution function (5.2) over the region $|x|/\sigma < c$. We should newly evaluate the probability of the type-1 error because the nonzero g case is included in contrast to the previous test in Chap. 5 We can calculate the probability by integrating the distribution function (5.2) over the region $|x|/\sigma > c$. Namely, the probability of the type-1 error in the weak measurement is given by

$$E_1(g_1, |\langle \hat{A} \rangle_w|^2) := 1 - E_2(g_1, |\langle \hat{A} \rangle_w|^2). \tag{D.4}$$

For the conventional measurement, we can substantially treat the probabilities of the type-1 and type-2 errors by substituting $|\langle \hat{A} \rangle_w|^2 = 1$ into Eqs. (D.4) and (D.3), respectively. For convenience, we uses the notations

$$h(g) = \mathrm{erf} \left[\frac{c\sigma - g}{\sqrt{2\sigma^2}} \right] + \mathrm{erf} \left[\frac{c\sigma + g}{\sqrt{2\sigma^2}} \right] \tag{D.5}$$

$$s := 1 + |\langle \hat{A} \rangle_w|^2, \tag{D.6}$$

$$t := 1 - |\langle \hat{A} \rangle_w|^2. \tag{D.7}$$

Hereafter, we attempt to calculate the condition for the weak value which satisfies the inequality

$$rE_1(g_1, \langle \hat{A} \rangle_w) + (1 - r)E_2(g_2, \langle \hat{A} \rangle_w) \leq rE_1(g_1, 1) + (1 - r)E_2(g_2, 1). \tag{D.8}$$

To solve this problem, we consider the function

$$\begin{aligned} \zeta[\langle \hat{A} \rangle_w] &= rE_1(g_1, \langle \hat{A} \rangle_w) + (1 - r)E_2(g_2, \langle \hat{A} \rangle_w) - [rE_1(g_1, 1) + (1 - r)E_2(g_2, 1)] \\ &= r[E_1(g_1, \langle \hat{A} \rangle_w) - E_1(g_1, 1)] + (1 - r)[E_2(g_2, \langle \hat{A} \rangle_w) - E_2(g_2, 1)] \\ &= -r[E_2(g_1, \langle \hat{A} \rangle_w) - E_2(g_1, 1)] + (1 - r)[E_2(g_2, \langle \hat{A} \rangle_w) - E_2(g_2, 1)]. \end{aligned} \tag{D.9}$$

We find the weak value such as the solution of $\zeta(\langle \hat{A} \rangle_w) = 0$. Because we can easily see that $t = 0 \Leftrightarrow |\langle \hat{A} \rangle_w|^2 = 1$ is the trivial solution, hereafter we assume $t \neq 0$ to find another

solution. $2\zeta[\langle\hat{A}\rangle_w] = 0$ becomes

$$
\begin{aligned}
&- r\left[\frac{sh(g_1) + te^{-g_1^2/2\sigma^2}h(0)}{s + te^{-g_1^2/2\sigma^2}} - h(g_1)\right] + (1-r)\left[\frac{sh(g_2) + te^{-g_2^2/2\sigma^2}h(0)}{s + te^{-g_2^2/2\sigma^2}} - h(g_2)\right] = 0 \\
&\quad\Leftrightarrow -r[sh(g_1) + te^{-g_1^2/2\sigma^2}](s + te^{-g_2^2/2\sigma^2}) \\
&\qquad + (1-r)[sh(g_2) + te^{-g_2^2/2\sigma^2}h(0)(s + te^{-g_1^2/2\sigma^2})] \\
&\qquad\quad - [-rh(g_1) + (1-r)h(g_2)](s + te^{-g_1^2/2\sigma^2})(s + te^{-g_2^2/2\sigma^2}) = 0 \\
&\quad\Leftrightarrow -s[-re^{-g_1^2/2\sigma^2}\eta(g_1) + (1-r)e^{-g_2^2/2\sigma^2}\eta(g_2)] \\
&\qquad - te^{-(g_1^2+g_2^2)/2\sigma^2}[-r\eta(g_1) + (1-r)\eta(g_2)] = 0,
\end{aligned}
\tag{D.10}
$$

where $\eta(g) := h(g) - h(0)^2$. From the equation, we can find the weak value of the solution candidate as

$$
\begin{aligned}
&|\langle\hat{A}\rangle_w|^2 \\
&= \frac{[-re^{-g_1^2/2\sigma^2}\eta(g_1) + (1-r)e^{-g_1^2/2\sigma^2}\eta(g_2)] + e^{-(g_1^2+g_2^2)/2\sigma^2}[-r\eta(g_1) + (1-r)\eta(g_2)]}{-[-re^{-g_1^2/2\sigma^2}\eta(g_1) + (1-r)e^{-g_1^2/2\sigma^2}\eta(g_2)] + e^{-(g_1^2+g_2^2)/2\sigma^2}[-r\eta(g_1) + (1-r)\eta(g_2)]} \\
&= \frac{-r\eta(g_1)(1 + e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1 + e^{g_1^2/2\sigma^2})}{-r\eta(g_1)(1 - e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1 - e^{g_1^2/2\sigma^2})} := A_s.
\end{aligned}
\tag{D.11}
$$

A_s can be a solution of $\zeta = 0$, if it is a positive real value. Especially, the solution such as $|\langle\hat{A}\rangle_w| > 1$ is expedient. The denominator of A_s can be not only positive but also negative. At first, we consider the case of the positive denominator, which provides the inequality

$$
\begin{aligned}
&- r\eta(g_1)(1 - e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1 - e^{g_1^2/2\sigma^2}) > 0 \\
&\Leftrightarrow 0 \le r < \frac{\eta(g_2)(1 - e^{g_1^2/2\sigma^2})}{\eta(g_1)(1 - e^{g_2^2/2\sigma^2}) + \eta(g_2)(1 - e^{g_1^2/2\sigma^2})} =: r_1.
\end{aligned}
\tag{D.12}
$$

And the condition for r to obtain the solution such as $|\langle\hat{A}\rangle_w| > 1$ is given by

$$
\begin{aligned}
A_s > 1 &\Leftrightarrow -r\eta(g_1)(1 + e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1 + e^{g_1^2/2\sigma^2}) \\
&\qquad > -r\eta(g_1)(1 - e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1 - e^{g_1^2/2\sigma^2}) \\
&\Leftrightarrow r_2 := \frac{\eta(g_2)e^{g_1^2/2\sigma^2}}{\eta(g_1)e^{g_2^2/2\sigma^2} + \eta(g_2)e^{g_1^2/2\sigma^2}} < r \le 1.
\end{aligned}
\tag{D.13}
$$

[2] $\eta(g)$ satisfies $\eta(0) = 0$ and $-1 \le \eta(g_2) < \eta(g_1) \le 0$. $\eta(g_1)$ and $\eta(g_2)$ are zero simultaneously only when $c \to \infty$, which gives $\zeta = 0$

However the magnitude relation between r_1 and r_2 is $r_2 > r_1$. There is no r region such that the denominator of A_s become positive and the solution of $\zeta = 0$ which satisfies $|\langle\hat{A}\rangle_w| > 1$ at the same time.

Next, we consider the case that the denominator of A_s is negative, which gives the inequality

$$-r\eta(g_1)(1-e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1-e^{g_1^2/2\sigma^2}) < 0 \Leftrightarrow r_1 < r \leq 1. \tag{D.14}$$

And the condition for r to obtain the solution such as $|\langle\hat{A}\rangle_w| > 1$ is

$$\begin{aligned} A_s > 1 \Leftrightarrow & -r\eta(g_1)(1+e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1+e^{g_1^2/2\sigma^2}) \\ & < -r\eta(g_1)(1-e^{g_2^2/2\sigma^2}) + (1-r)\eta(g_2)(1-e^{g_1^2/2\sigma^2}) \\ \Leftrightarrow & 0 \leq r < r_2. \end{aligned} \tag{D.15}$$

Thus we conclude that the region of r for giving the solution of $\zeta = 0$ such as $|\langle\hat{A}\rangle_w| > 1$ is $r_1 < r < r_2$.

Here we evaluate the limiting value of $\zeta[\langle\hat{A}\rangle_w]$ at $|\langle\hat{A}\rangle_w| \to \infty$. Using

$$\lim_{|\langle\hat{A}\rangle_w|\to\infty} t/s = \lim_{|\langle\hat{A}\rangle_w|\to\infty} (1-|\langle\hat{A}\rangle_w|)/(1+|\langle\hat{A}\rangle_w|) = -1, \tag{D.16}$$

the limiting value can be calculated as

$$\begin{aligned} & \lim_{|\langle\hat{A}\rangle_w|\to\infty} \zeta[\langle\hat{A}\rangle_w] \\ & = \lim_{|\langle\hat{A}\rangle_w|\to\infty} -\frac{r}{2}\left[\frac{h(g_1)+(t/s)e^{-g_1^2/2\sigma^2}h(0)}{1+(t/s)e^{-g_1^2/2\sigma^2}} - h(g_1)\right] \\ & \qquad + \frac{1-r}{2}\left[\frac{h(g_2)+(t/s)e^{-g_2^2/2\sigma^2}h(0)}{1+(t/s)e^{-g_2^2/2\sigma^2}} - h(g_2)\right] \\ & = -\frac{r}{2}\left[\frac{h(g_1)-e^{-g_1^2/2\sigma^2}h(0)}{1-e^{-g_1^2/2\sigma^2}} - h(g_1)\right] + \frac{1-r}{2}\left[\frac{h(g_2)-e^{-g_2^2/2\sigma^2}h(0)}{1-e^{-g_2^2/2\sigma^2}} - h(g_2)\right] \\ & = r\frac{\eta(g_1)}{2(1-e^{g_1^2/2\sigma^2})} - (1-r)\frac{\eta(g_2)}{2(1-e^{g_2^2/2\sigma^2})} = \frac{\eta(g_2)}{2(1-e^{g_2^2/2\sigma^2})}\left(\frac{r}{r_1}-1\right). \end{aligned} \tag{D.17}$$

Because $\eta(g_2)/(1-e^{g_2^2/2\sigma^2})$ is always positive, we find that r such as $r_1 < r \leq 1$ gives the positive limiting value of ζ.

Totally, we conclude that the the weighting factor r and the absolute value of the weak value $|\langle\hat{A}\rangle_w|$determine the case that the weak measurement is advantageous. We note the consequence as following.

- When $0 \leq r \leq r_1$, the weak measurement with the weak value such as $|\langle\hat{A}\rangle_w| > 1$ is superior to the conventional one.
- When $r_1 < r < r_2$, the weak measurement is advantageous with the weak value such as $1 < |\langle\hat{A}\rangle_w| < \sqrt{A_s}$.
- When $r_2 \leq r \leq 1$, the weak measurement has no advantage in any weak value.

Thus, we find that the weak measurement tends to be significant for the experiment such that the type-2 error, i.e., missing the interaction is more serious than the type-1 error, i.e., the false alarm. This conclusion is consistent with that given in the previous chapter, where we have evaluated the probability of the type-2 error of the two measurements.

D.3 Hypothesis Testing including data loss by post-selection

In Chap. 5, we assume that the data loss due to the failure of the postselection is neglected. However, some researchers strongly argue that the data loss by the postselection can cause the technical demerit of the WVA. Here, we discuss the data loss problem with the makeshift test function and we derive the optimal case with the Lagrange multiplier.

To consider the no data case which results from the failure of the posetselection, we propose the different test function from Eq. (5.4) as

$$d(x) = \begin{cases} 0 & \text{if } (f \text{ and } |x|/\sigma < c_f) \text{ or } (\bar{f} \text{ and } |x|/\sigma < c_{\bar{f}}), \\ 1 & \text{if } (f \text{ and } |x|/\sigma > c_f) \text{ or } (\bar{f} \text{ and } |x|/\sigma > c_{\bar{f}}), \\ r & \text{otherwise,} \end{cases} \tag{D.18}$$

where f and $\bar{f}$ represent success and failure of the postselection, respectively. With this test function, we can include the result of the postselection and the measurement result x. The critical c_f and $c_{\bar{f}}$ are different in response to the results of the postselection.

Next, we evaluate the probabilities of the two types of the error. The probability of the type-1 error is

$$\begin{aligned}
&\Pr[\mathcal{E}_1] \\
&= \Pr[d=1|g=0] \\
&= \Pr[f, |x|/\sigma > c_f | g=0] + \Pr[\bar{f}, |x|/\sigma > c_{\bar{f}} | g=0] \\
&= 1 - \left(\mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right] |\langle f|i\rangle|^2 + \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right] |\langle \bar{f}|i\rangle|^2 \right),
\end{aligned} \tag{D.19}$$

and the probability of the type-2 error is given as

$$\begin{aligned}
&\Pr[\mathcal{E}_2] \\
&= \Pr[d=0|g\neq 1] \\
&= \Pr[f, |x|/\sigma < c_f | g\neq 0] + \Pr[\bar{f}, |x|/\sigma < c_{\bar{f}} | g\neq 0] \\
&= \frac{1}{4}(|\langle f|i\rangle|^2 + |\langle f|\hat{A}|i\rangle|^2) \left(\mathrm{erf}\left[\frac{c_f\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c_f\sigma + g}{\sqrt{2\sigma^2}}\right] \right) \\
&\quad + \frac{1}{4}(|\langle \bar{f}|i\rangle|^2 + |\langle \bar{f}|\hat{A}|i\rangle|^2) \left(\mathrm{erf}\left[\frac{c_{\bar{f}}\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c_{\bar{f}}\sigma + g}{\sqrt{2\sigma^2}}\right] \right) \\
&\quad + \frac{1}{2}(|\langle f|i\rangle|^2 - |\langle f|\hat{A}|i\rangle|^2) e^{-\frac{g^2}{2\sigma^2}} \mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right] \\
&\quad + \frac{1}{2}(|\langle \bar{f}|i\rangle|^2 - |\langle \bar{f}|\hat{A}|i\rangle|^2) e^{-\frac{g^2}{2\sigma^2}} \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right].
\end{aligned} \tag{D.20}$$

Here we tentatively set $c_f = c_{\bar{f}} (= c)$ and the probability of the type-1 error becomes

$$\Pr[\mathcal{E}_1] = 1 - \mathrm{erf}\left[\frac{c}{\sqrt{2}}\right], \tag{D.21}$$

and the probability of the type-2 error becomes

$$\Pr[\mathcal{E}_2] = \frac{1}{2}\left(\mathrm{erf}\left[\frac{c\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c\sigma + g}{\sqrt{2\sigma^2}}\right] \right). \tag{D.22}$$

Both the probabilities are independent on the postselection results $|f\rangle, |\bar{f}\rangle$. Also, we find that these probabilities are same as the ones given by the conventional measurement [Eqs. (5.6) and (5.7)]. We can virtually treat the conventional measurement case by setting $c_f = c_{\bar{f}}$ with the revised decision function (D.18).

Next we examine the errors of the weak measurement in the decision function (D.18). First, we consider the case of that we cannot obtain a measurement result due to post-selection failure. In such a case, we cannot distinguish whether or not the interaction

is presence. For the moment, we simply presume that the interaction will be absence. Because we are interested in the detection of the interaction with obtained data in an usual experiment, the decision with no data is worthless. Setting $c_{\bar{f}} = \infty$, we can handle this situation in our decision function for convenience. With $c_{\bar{f}} = \infty$, the alternative hypothesis is always rejected by the test (D.18) when the postselection fails. Hence, the decision function (D.18) would cover the case with or without postselection including the data loss by failure of the postselection.

Here we have to pay attention the following matters. There will be an experimental problem that we want to detect vanishment of an interaction which usually exists. In such experiment, the treatment of $c_{\bar{f}}$ as stated above is inconsequent. Also we note that if the c_f and $c_{\bar{f}}$ 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 $c_{\bar{f}}$. 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 (D.18). To optimize the probability of the type-2 error $\Pr[\mathcal{E}_2]$ while keeping the probability of the type-1 error $\Pr[\mathcal{E}_1]$ at the significance level α which is an arbitrary constant, we fix the Lagrangian as

$$\begin{aligned}
&\mathcal{L}(p_1,\ p_2,\ c_f,\ c_{\bar{f}},\ \lambda) \\
&\quad = \Pr[\mathcal{E}_2] + \lambda(\Pr[\mathcal{E}_1] - \alpha) \\
&\quad = \frac{1}{4}\left[(p_1+p_2)\left(\mathrm{erf}\left[\frac{c_f\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c_f\sigma + g}{\sqrt{2\sigma^2}}\right]\right)\right. \\
&\qquad\quad + (2-p_1-p_2)\left(\mathrm{erf}\left[\frac{c_{\bar{f}}\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c_{\bar{f}}\sigma + g}{\sqrt{2\sigma^2}}\right]\right) \\
&\qquad\quad \left. +2(p_1-p_2)\left(\mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right]\right)e^{-\frac{g^2}{2\sigma^2}}\right] \\
&\qquad + \lambda\left[p_1\left(1 - \mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right]\right) + (1-p_1)\left(1 - \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right]\right) - \alpha\right],
\end{aligned} \tag{D.23}$$

where λ is the Lagrange multiplier and the constraint condition comes from the standard strategy of the hypothesis testing as stated in Sec. 4.2. Here we have denoted $p_1 := |\langle f|i\rangle|^2$ and $p_2 := |\langle f|\hat{A}|i\rangle|^2$ for convenience. We note that $\hat{A}^2 = 1$ and $0 < p_1, p_2 < 1$. Varying the Lagrangian $\mathcal{L}$ with respect to λ, the constraint condition reappears as

$$0 = \frac{\partial\mathcal{L}}{\partial\lambda} = p_1\left(1 - \mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right]\right) + (1-p_1)\left(1 - \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right]\right) - \alpha. \tag{D.24}$$

Next, varying the Lagrangian $\mathcal{L}$ with respect to p_1 and p_2, we obtain the equations,

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}}{\partial p_1} \\ &= \frac{1}{4}\left[\mathrm{erf}\left[\frac{c_f\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c_f\sigma + g}{\sqrt{2\sigma^2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}\sigma - g}{\sqrt{2\sigma^2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}\sigma + g}{\sqrt{2\sigma^2}}\right]\right. \\ &\qquad \left. +2\left(\mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right]\right)e^{-\frac{g^2}{2\sigma^2}}\right] + \lambda\left(-\mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right] + \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right]\right)\end{aligned} \tag{D.25}$$

and

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}}{\partial p_2} \\ &= \frac{1}{4}\left[\mathrm{erf}\left[\frac{c_f\sigma - g}{\sqrt{2\sigma^2}}\right] + \mathrm{erf}\left[\frac{c_f\sigma + g}{\sqrt{2\sigma^2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}\sigma - g}{\sqrt{2\sigma^2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}\sigma + g}{\sqrt{2\sigma^2}}\right]\right. \\ &\qquad \left. -2\left(\mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right]\right)e^{-\frac{g^2}{2\sigma^2}}\right],\end{aligned} \tag{D.26}$$

which lead the relation

$$0 = \frac{\partial \mathcal{L}}{\partial p_1} - \frac{\partial \mathcal{L}}{\partial p_2} = \left(\mathrm{erf}\left[\frac{c_f}{\sqrt{2}}\right] - \mathrm{erf}\left[\frac{c_{\bar{f}}}{\sqrt{2}}\right]\right)(e^{-\frac{g^2}{2\sigma^2}} - \lambda). \tag{D.27}$$

So, we require either or both of $\lambda = e^{-g^2/2\sigma^2}$ and $c_f = c_{\bar{f}}$ for Eq. (D.27) . Varying $\mathcal{L}$ with respect to c_f and $c_{\bar{f}}$, we obtain the equations

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}}{\partial c_f} \\ &= \frac{1}{4}\sqrt{\frac{2}{\pi}}\left[(p_1 + p_2)(e^{-\frac{(c_f\sigma - g)^2}{2\sigma^2}} + e^{-\frac{(c_f\sigma + g)^2}{2\sigma^2}}) + 2(p_1 - p_2 - 2\lambda p_1 e^{\frac{g^2}{2\sigma^2}})e^{-\frac{c_f^2\sigma^2 + g^2}{2\sigma^2}}\right]\end{aligned} \tag{D.28}$$

and

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}}{\partial c_{\bar{f}}} \\ &= \frac{1}{4}\sqrt{\frac{2}{\pi}}\left[(2 - p_1 - p_2)\big(e^{-\frac{(c_{\bar{f}}\sigma - g)^2}{2\sigma^2}} + e^{-\frac{(c_{\bar{f}}\sigma + g)^2}{2\sigma^2}}\big)\right. \\ &\qquad \left. +2\big(-p_1 + p_2 - 2\lambda(1 - p_1)e^{\frac{g^2}{2\sigma^2}}\big)e^{-\frac{c_{\bar{f}}^2\sigma^2 + g^2}{2\sigma^2}}\right].\end{aligned} \tag{D.29}$$

Here, we consider the case $c_f = c_{\bar{f}} (=: c)$, where Eqs. (D.25) and (D.26) are fulfilled. The constraint constant (D.24) becomes

$$0 = \left(1 - \mathrm{erf}\left[\frac{c}{\sqrt{2}}\right]\right) - \alpha. \tag{D.30}$$

Because the α is a constant, the c is fixed. Eqs. (D.28) and (D.29) give

$$0 = \frac{e^{-\frac{c^2}{2}}}{4}\sqrt{\frac{2}{\pi}}\left[p_1\left\{e^{-\frac{g^2}{2\sigma^2}}(e^{\frac{cg}{2\sigma}} + e^{-\frac{cg}{2\sigma}})^2 - 4\lambda\right\} + p_2 e^{-\frac{g^2}{2\sigma^2}}(e^{\frac{cg}{2\sigma}} - e^{-\frac{cg}{2\sigma}})^2\right] \tag{D.31}$$

and

$$\begin{aligned} 0 = \frac{e^{-\frac{c^2}{2}}}{4}\sqrt{\frac{2}{\pi}}\Big[&2\big\{e^{-\frac{g^2}{2\sigma^2}}(e^{\frac{cg}{\sigma}} + e^{-\frac{cg}{\sigma}}) - 2\lambda\big\} \\ &-p_1\big\{e^{-\frac{g^2}{2\sigma^2}}(e^{\frac{cg}{2\sigma}} + e^{-\frac{cg}{2\sigma}})^2 - 4\lambda\big\} - p_2 e^{-\frac{g^2}{2\sigma^2}}(e^{\frac{cg}{2\sigma}} - e^{-\frac{cg}{2\sigma}})^2\Big], \end{aligned} \tag{D.32}$$

respectively. The sum of Eqs. (D.31) and (D.32) is

$$0 = \frac{1}{2}\sqrt{\frac{2}{\pi}}e^{-\frac{c^2}{2}}\big\{e^{-\frac{g^2}{2\sigma^2}}(e^{\frac{cg}{\sigma}} + e^{-\frac{cg}{\sigma}}) - 2\lambda\big\}. \tag{D.33}$$

Thus, we obtain

$$\lambda = \frac{1}{2}e^{-\frac{g^2}{2\sigma^2}}(e^{\frac{cg}{\sigma}} + e^{-\frac{cg}{\sigma}}). \tag{D.34}$$

Substituting $c_f = c_{\bar{f}} = c$ and Eq. (D.34) into Eqs. (D.28) and (D.29), we have

$$0 = \frac{e^{-\frac{c^2\sigma^2+g^2}{2\sigma^2}}}{4}\sqrt{\frac{2}{\pi}}(p_1 - p_2)\big(e^{\frac{cg}{2\sigma}} - e^{-\frac{cg}{2\sigma}}\big)^2. \tag{D.35}$$

Therefore we can find that $c_f = c_{\bar{f}} = c = 0$ or $p_1 = p_2$ is needed. If $c = 0$, we obtain $\alpha = 1$ and $\lambda = e^{-g^2/2\sigma^2}$ from Eqs. (D.30) and (D.34). Because the significance level α is not always 1, the $c = 0$ is not consistent. Thus $c_f = c_{\bar{f}} \neq 0$ and $p_1 = p_2$ is a solution of the Lagrange multiplier problem.

Next we discuss the case $\lambda = e^{-g^2/2\sigma^2}$ which is derived from Eq. (D.27). Because we have already seen the case that $\lambda = e^{-g^2/2\sigma^2}$ and $c_f = c_{\bar{f}}$ are satisfied simultaneously, hereafter we assume $c_f \neq c_{\bar{f}}$. Substituting $\lambda = e^{-g^2/2\sigma^2}$ into Eqs. (D.28) and (D.29), we

obtain

$$0 = \frac{1}{4}\sqrt{\frac{2}{\pi}}(p_1 + p_2)\left(e^{\frac{c_f g}{2\sigma}} - e^{-\frac{c_f g}{2\sigma}}\right)^2 e^{-\frac{c_f^2\sigma^2 + g^2}{2\sigma^2}}, \tag{D.36}$$

and

$$0 = \frac{1}{4}\sqrt{\frac{2}{\pi}}(2 - p_1 - p_2)\left(e^{\frac{c_{\bar{f}} g}{2\sigma}} - e^{-\frac{c_{\bar{f}} g}{2\sigma}}\right)^2 e^{-\frac{c_{\bar{f}}^2\sigma^2 + g^2}{2\sigma^2}}. \tag{D.37}$$

From these equations, 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 $c_{\bar{f}} = 0$ and $p_1 = p_2 = 0$. In both cases, the constrain condition (D.24) becomes $0 = 1 - \alpha$. As stated in above, the significance point α is not necessarily 1. So, the condition $\lambda = e^{-g^2/2\sigma^2}$ is not proper.

Eventually, we conclude the solution is $c_f = c_{\bar{f}} \neq 0$ and $p_1 = p_2$. From $p_1 = p_2$, we derive

$$0 = |\langle f|i\rangle|^2 - |\langle f|\hat{A}|i\rangle|^2 = \langle i|(|f\rangle\langle f| - \hat{A}|f\rangle\langle f|\hat{A})|i\rangle, \;\; \therefore \;\; \pm|f\rangle = \hat{A}|f\rangle \tag{D.38}$$

or

$$0 = \left|\langle f|i\rangle\right|^2 - \left|\langle f|\hat{A}|i\rangle\right|^2 = \langle f|(|i\rangle\langle i| - \hat{A}|i\rangle\langle i|\hat{A})|f\rangle, \;\; \therefore \;\; \pm|i\rangle = \hat{A}|i\rangle. \tag{D.39}$$

$p_1 = p_2$ implies that the preselected state $|i\rangle$ or the postselected state $|f\rangle$ equals to an eigenstate of $\hat{A}$. The case $c_f = c_{\bar{f}}$ corresponds to the conventional measurement which does not have postselection, which intend the result of the postselection has nothing to do with the test under the solution of the variational problem. The state of the postselection $|f\rangle$ is not essential. Consequently, the derived condition is that the preselected state is an eigenstate of the measured observable $\hat{A}$ and that we do not postselect.

Again we have to care that this concussion is deduced from a makeshift decision function (D.18) as we noted the some defect in above. We also note that the Lagrange multiplier method gives the stationary point to the utmost, and they might not be the minimum.

References

[1] M. Ozawa, Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement, Phys. Rev. A **67**, 042105 (2003).

[2] Y. Watanabe, T. Sagawa, and M. Ueda, Uncertainty relation revisited from quantum estimation theory, Phys. Rev. A **84**, 042121 (2011).

[3] C. Branciard, Error-tradeoff and error-disturbance relations for incompatible quantum measurements, Proc. Natl. Acad. Sci. U.S.A. **110**, 6742 (2013).

[4] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, Zeitschrift Für Phys. **43**, 172 (1927).

[5] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, Experimental demonstration of a universally valid error - disturbance uncertainty relation in spin measurements, Nat. Phys. **8**, 185 (2012).

[6] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, Experimental Test of Error-Disturbance Uncertainty Relations by Weak Measurement, Phys. Rev. Lett. **112**, 020402 (2014).

[7] B. P. Abbott, *et al.*, LIGO: the Laser Interferometer Gravitational-Wave Observatory, Rep. Prog. Phys. **72**, 076901 (2009).

[8] K. Somiya, Detector configuration of KAGRA-the Japanese cryogenic gravitational-wave detector, Class. Quantum Gravity **29**, 124007 (2012).

[9] H. J. Kimble, Y. Levin, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin, Conversion of conventional gravitational-wave interferometers into quantum nondemolition interferometers by modifying their input and/or output optics, Phys. Rev. D **65**, 022002 (2001).

[10] V. B. Braginsky, M. L. Gorodetsky, F. Y. Khalili, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin, Noise in gravitational-wave detectors and other classical-force measurements is not influenced by test-mass quantization, Phys. Rev. D **67**, 082001 (2003).

[11] T. J. Kippenberg and K. J. Vahala, Cavity Optomechanics: Back-Action at the Mesoscale, Science **321**, 1172 (2008).

[12] S. L. Danilishin and F. Y. Khalili, Quantum Measurement Theory in Gravitational-Wave Detectors, Living Rev. Relativ. **15**, 5 (2012).

[13] Y. Chen, Macroscopic quantum mechanics: theory and experimental concepts of optomechanics, J. Phys. B: At. Mol. Opt. Phys. **46**, 104001 (2013).

[14] K. J. Resch, Amplifying a tiny optical effect, Science **319**, 733 (2008).

[15] S. Popescu, Weak measurements just got stronger, Physics, **2**, 32 (2009).

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

[17] 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).

[18] 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).

[19] 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).

[20] 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).

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

[22] 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).

[23] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Optimizing the signal-to-noise ratio of a beam-deflection measurement with interferometric weak values, Phys. Rev. A **80**, 041803(R) (2009).

[24] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, Continuous phase amplification with a Sagnac interferometer, Phys. Rev. A **82**, 011802(R) (2010).

[25] M. D. Turner, C. A. Hagedorn, S. Schlamminger, and J. H. Gundlach, Picoradian deflection measurement with an interferometric quasi-autocollimator using weak value amplification, Opt. Lett. **36**, 1479 (2011).

[26] J. M. Hogan, J. Hammer, S.-W. Chiow, S. Dickerson, D. M. S. Johnson, T. Kovachy, A. Sugarbaker, and M. A. Kasevich, Precision angle sensor using an optical lever inside a Sagnac interferometer, Opt. Lett. **36**, 1698 (2011).

[27] M. Pfeifer and P. Fischer, Weak value amplified optical activity measurements Opt. Express **19**, 16508 (2011).

[28] G. I. Viza, J. Martines-Rincon, G. A. Howland, H. Frostig, I. Shomroni, B. Dayan, and J. C. Howell, Weak-values technique for velocity measurements, Opt. Lett. **38**, 2949-2952 (2013).

[29] G. Jayaswal, G. Mistura, and M. Merano, Observation of the Imbert - Fedorov effect via weak value amplification, Opt. Lett. **39**, 2266 (2014).

[30] O. S. Maga ñ a-Loaiza, M. Mirhosseini, B. Rodenburg, and R. W. Boyd, Amplification of Angular Rotations Using Weak Measurements, Phys. Rev. Lett. **112**, 200401 (2014).

[31] S. Wu and Y. Li, Weak measurements beyond the Aharonov-Albert-Vaidman formalism, Phys. Rev. A **83**, 052106 (2011).

[32] X. Zhu, Y. Zhang, S. Pang, C. Qiao, Q. Liu, and S. Wu, Quantum measurements with preselection and postselection, Phys. Rev. A **84**, 052111 (2011).

[33] T. Koike and S. Tanaka, Limits on amplification by Aharonov-Albert-Vaidman weak measurement, Phys. Rev. A **84**, 062106 (2011).

[34] K. Nakamura, A. Nishizawa, and M.-K. Fujimoto, Evaluation of weak measurements to all orders, Phys. Rev. A **85**, 012113 (2012).

[35] S. Pang, T. A. Brun, S. Wu, and Z.-B. Chen, Amplification limit of weak measurements: A variational approach, Phys. Rev. A 90, 012108 (2014).

[36] A. Nishizawa, K. Nakamura, and M.-K. Fujimoto, Weak-value amplification in a shot-noise-limited interferometer, Phys. Rev. A **85**, 062108 (2012).

[37] A. Nishizawa, Weak-value amplification beyond the standard quantum limit in position measurements, Phys. Rev. A **92**, 032123 (2015).

[38] 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).

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

[40] 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).

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

[42] C. Ferrie and J. Combes, Weak value amplification is suboptimal for estimation and detection, Phys. Rev. Lett. **112**, 040406 (2014).

[43] G. C. Knee and E. M. Gauger, When amplification with weak values fails to suppress technical noise. Phys. Rev. X **4**, 011032 (2014).

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

[45] A. Takemura, *Modern Mathematical Statistics* (Sobunsha, Tokyo, 1991, in Japanese).

[46] J. Dressel, K. Lyons, A. N. Jordan, T. M. Graham, and P. G. Kwiat, Strengthening weak-value amplification with recycled photons, Phys. Rev. A **88**, 023821 (2013).

[47] K. Lyons, J. Dressel, A. N. Jordan, J. C. Howell, and P. G. Kwiat, Power-Recycled Weak-Value-Based Metrology, Phys. Rev. Lett. **114**, 170801 (2015).

[48] S. Pang, J. Dressel, and T. A. Brun, Entanglement-Assisted Weak Value Amplification, Phys. Rev. Lett. **113**, 030401 (2014).

[49] S. Pang and T. A. Brun, Suppressing technical noise in weak measurements by entanglement, Phys. Rev. A **92**, 012120 (2015).

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

[51] Y. Kedem, Comment on "Weak value amplification is suboptimal for estimation and detection", arXiv:1402.1352.

[52] C. Ferrie and J. Combes, Reply to comments on "Weak value amplification is suboptimal for estimation and detection", arXiv:1402.2954.

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

[54] C. R. Rao, *Linear Statistical Inference and its Applications* 2nd ed. (Wiley, New York, 1973).

[55] J. C. Kiefer, *Introduction to statistical inference* (Springer, New York, 1987).

[56] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. (Springer, NewYork, 1985), pp. 20-21.

[57] J. O. Berger and M. Delampady, Testing precise hypotheses, Stat. Sci. **2**, 317 (1987).

[58] Y. Susa, Y. Shikano, and A. Hosoya, Optimal probe wave function of weak-value amplification, Phys. Rev. A **85**, 052110 (2012).

[59] Y. Susa, Y. Shikano, and A. Hosoya, Reply to "Comment on 'Optimal probe wave function of weak-value amplification' ", Phys. Rev. A **87**, 046102 (2013).

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

[61] Y. Susa, Physical description of statistical hypothesis testing for a weak-value-amplification experiment using a birefringent crystal, Phys. Rev. A **92**, 022118 (2015).

[62] A. D. Lorenzo, Comment on "Optimal probe wave function of weak-value amplification", Phys. Rev. A **87**, 046101 (2013).

[63] J. Neyman and E. S. Pearson, On the problem of the most efficient tests of statistical hypotheses, Philos. Trans. R. Soc. A: Math., Phys. Eng. Sci. **231**, 289 (1933).

[64] E. L. Lehmann, *Testing Statistical Hypotheses* (Wiley, New York, 1959).

[65] T. Ando, Y. Ohtake, N. Matsumoto, T. Inoue, and N. Fukuchi, Mode purities of Laguerre-Gaussian beams generated via complex-amplitude modulation using phase-only spatial light modulators, Opt. Lett. **34**, 34 (2009).

[66] H. Kobayashi, G. Puentes, and Y. Shikano, Extracting joint weak values from two-dimensional spatial displacements, Phys. Rev. A **86**, 053805 (2012).

[67] H. Kobayashi, G. Puentes, and Y. Shikano, Stereographical visualization of a polarization state using weak measurements with an optical-vortex beam, Phys. Rev. A **86**, 053816 (2014).

[68] Y. Turek, H. Kobayashi, T. Akutsu, C.-P. Sun, and Y. Shikano, Post-selected von Neumann measurement with Hermite-Gaussian and Laguerre-Gaussian pointer states, New J. Phys. **17**, 083029 (2015).

[69] H. Jeffreys, *Theory of Probability*, 3rd ed. (Oxford Univ. Press, Oxford, 1961).

[70] R. E. Kass and A. E. Raftery, Bayes factors, J. Am. Stat. Assoc. **90**, 773 (1995).

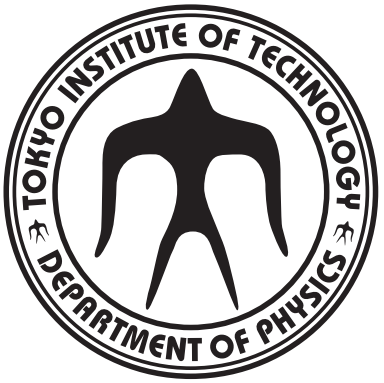
[71] A. C. Elitzur and L. Vaidman, Quantum mechanical interaction-free measurements, Found. Phys. **23**, 987 (1993).

[72] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M. A. Kasevich, Interaction-free measurement, Phys. Rev. Lett. **74**, 4763 (1995).

[73] O. Zilberberg, A. Romito, and Y. Gefen, Null weak values in multi-level systems, Phys. Scr. **T151**, 014014 (2012).

[74] O. Zilberberg, A. Romito, D. J. Starling, G. a. Howland, C. J. Broadbent, J. C. Howell, and Y. Gefen, Null Values and Quantum State Discrìmination, Phys. Rev. Lett. **110**, 170405 (2013).

[75] O. Zilberberg, A. Romito, and Y. Gefen, Standard and Null Weak Values, in *Quantum Theory: A Two-Time Success Story*, edited by D. C. Struppa and J. M. Tollaksen (Springer Milan, Milano, 2014), pp. 377-387.


TOKYO INSTITUTE OF TECHNOLOGY
DEPARTMENT OF PHYSICS