

Doctoral Dissertation

博士論文

On the Possibility of Application of Weak
Value Amplification to Precision Measurement

(精密測定への弱値増幅の応用可能性について)

A Dissertation Submitted for the Degree of Doctor of Philosophy
December 2020

令和2年12月博士（理学）申請

Department of Physics, Graduate School of Science,
The University of Tokyo

東京大学大学院理学系研究科物理学専攻

Yuichiro Mori

森 雄一朗

Abstract

The novel notion of a physical quantity called ‘weak value’, which was proposed by Aharonov *et al.* in 1988, has attracted much attention in the fields of quantum foundations and quantum measurement. In particular, the fact that through proper state selections the weak value can be made arbitrarily large, exceeding the conventional bound set by the maximum eigenvalue of the physical observable, has been widely regarded as a promising source for precision measurement. Indeed, in recent years we have seen increasingly many reports of successful applications of the weak value in photonic systems with various instances of observables amplified effectively. However, it has also been recognized that there are two major obstacles in realizing this weak value amplification (WVA). First, it is unclear under what conditions the WVA becomes effective due to the conflict between the ambiguity caused by the state selections and the amplification effect. Second, the required state selections may not be available depending on the systems one is dealing with. In order to expand the applicability of the WVA, one must solve these problems properly.

In this dissertation, we first review the basics of precision measurement and the WVA and then present our method of evaluation for the WVA considering the measurement uncertainty including the ambiguity of the WVA. We applied our method of evaluation for the two celebrated experiments to confirm the basic validity of the results of the two experiments. In addition, as a new application of the WVA, we consider the time evolution of unstable particles to find that the decay time distribution is generally characterized by the weak value of the Hamiltonian and that the state selections can indeed be performed by measuring the particles produced during the decay. Our result indicates that the WVA may also be applicable to the system of unstable particles.

Contents

1	Introduction	3
2	Metrology and Precision Measurement	7
2.1	Measurement and theoretical uncertainties	7
2.2	Random uncertainty and systematic uncertainty	8
2.3	Direct and indirect measurement in metrology	8
2.4	Parameter estimation and its uncertainty	10
2.4.1	The maximum likelihood estimation method	10
2.5	Summary	12
3	Weak Value Amplification	13
3.1	Indirect measurement in quantum measurement	13
3.1.1	The von Neumann measurement model	14
3.1.2	Weakness of the measurement	15
3.1.3	Measurement on the meter system	16
3.2	Aharonov's weak measurement	16
3.3	Weak value amplification (WVA)	19
3.3.1	On the possible weak value	19
3.3.2	'Tail' can be more affected from the interaction	20
3.3.3	Other schemes for deriving the weak value	21
3.3.4	Preceding Experiments	22
3.4	Controversial points of the WVA	25
3.4.1	Issues involving the decreasing of the number	26
3.4.2	Dependence of the meter state	28
3.5	Technical points in actual experiments	28
3.5.1	Setting of target system and meter system in experiments using photons	29
3.5.2	Points to consider common to optical experiments	31
3.5.3	Examples of state selection in the target system	33
4	Effects on the WVA	37
4.1	Measurement Uncertainty	37
4.2	Calculation model	39
4.2.1	Model uncertainty	40
4.2.2	Approximation uncertainty	40
4.2.3	Statistical uncertainty	42
4.2.4	Intractable uncertainty	44
4.3	WVA in individual experiments	45
4.3.1	SHEL experiment	45
4.3.2	USBD experiment	46

4.4	Parameter settings for simulation of each experiment	47
4.4.1	The number of statistics N	48
4.4.2	The intractable uncertainty δ_{int}	48
4.5	Results of our analysis	49
4.5.1	The SHEL experiment	49
4.5.2	The USBD experiment	50
4.6	Relevance to other studies	51
4.7	Summary and Discussions of this chapter	52
5	An application of WVA to high energy experiments	55
5.1	Weak Value in Particle Decay	55
5.1.1	Meter system like understanding of this type amplification	57
5.2	Dynamics of B meson and Time distributions	58
5.2.1	Hamiltonian	58
5.2.2	Time distribution without postselection	60
5.2.3	Time distribution with postselection	61
5.3	Weak value amplification	62
5.4	Process of postselection and related issues	64
5.5	Issues related to experimental method and systems	66
5.5.1	Extension of time distribution in the negative direction	67
5.5.2	Time measurement method	67
5.6	Uncertainty evaluation with Monte-Carlo simulation	69
5.6.1	Event simulation	69
5.6.2	Signal extraction and uncertainty in this analysis	71
5.7	Analysis result	73
5.8	Comments on the time distribution and Monte-Carlo simulation	74
5.8.1	On the derivation of the time distribution	74
5.8.2	On the validity of the results on the estimated uncertainty	75
6	Conclusion and Discussions	77
A	On the state of the photon and time evolution	83
A.1	Quantum-Classical Correspondence	83
A.2	Free propagation	84
A.3	Lens as a unitary operator	85
B	Some useful results	87
B.1	Momentum as the generator of translation	87
B.2	The relation between the eigenstates of the flavor and the Hamiltonian	87
B.3	Monotonicity of the Lower-bound Function	88
B.4	Uncorrelated intractable error	89
B.5	Calculation involving the Gaussian integrals	90

Chapter 1

Introduction

Measurement has been an essential element of science and technology, and modern society has greatly benefited from them. However, we are fully aware that the values obtained by measurements contain certain uncertainty in general. Although the cause of the error may be inferred, it is difficult to fully identify the uncertainty that occurred in the actual measurement, and in that sense, the discussion of the error in measurement is necessarily equivalent to a discussion of uncertainty in measurement. The uncertainty contained in the measurement results constrains our quantitative discussions. It is our great hope to reduce these constraints by reducing the uncertainty considerably. Accordingly, measurements with minimal uncertainty are desirable for us, and we refer to these measurements in this dissertation as precision measurements.

Precision measurements generally play an important role in science. An instructive example can be found in the relationship between the improvement of measurement accuracy for the electron g -factor and the development of physical theory. First, the fact that the g -factor is closer to 2 than 1 indicates that the behavior of the electron is relativistic. Next, the deviation of the g -factor from 2 was explained beautifully as the effect of quantum corrections [1, 2], which supports the correctness of the quantum electrodynamics that assume not only relativity but also quantum theory. However, later we found that the muon's g -factor deviates slightly from the theoretical calculations. This problem is called muon's $g - 2$ problem, which motivates various studies to improve the accuracy of measurements as well as theoretical calculation in the Standard Model, and the pursuit of new theories that go beyond it.

Meanwhile, quantum mechanics was built in the early part of the last century with a laborious effort of understanding the physics and experiments at that time. Quantum mechanics was applied to minute physical systems such as atoms [3] and nuclei from an early stage, and contributed to deepening the understanding of such minute physical systems. Besides, quantum mechanics gave a crucial impact on our understanding of measurement. In fact, when Heisenberg proposed the uncertainty principle from the early stages of quantum mechanics [4], he declared that the effect of measurement on the state is unavoidable. After that, the understanding of measurement based on quantum theory was greatly advanced [5, 6].

Quantum mechanics is a theory that calculates the probability of transition of states in a measurement performed for some prepared state. More specifically, we have the quantity called “amplitude”,

$$\langle \phi | \psi \rangle, \tag{1.1}$$

where $|\psi\rangle$ is the initial state and $|\phi\rangle$ is the state corresponding to the result of the measurement, and the absolute square of the amplitude gives the probability we want to find.

In that sense, quantum theory is a probability theory characterized by two states, $|\psi\rangle$ and $|\phi\rangle$.

In 1964, Aharonov, Bergmann, and Lebowitz reconsidered the formalism of quantum mechanics and proposed a new formalism in which the states $|\psi\rangle$ and $|\phi\rangle$ are treated equally [7]. This time-symmetric formalism of quantum mechanics raised a question as to what the expectation value of an observable \hat{A} is when both of the state $|\psi\rangle$ and $|\phi\rangle$ are specified, which are referred to as *preselected state* and *postselected state*, respectively. The expectation value which appears in these circumstances is therefore a conditional expectation value. The answer to the question is given by a novel quantity called a *weak value* [8]

$$A_w := \frac{\langle\phi|\hat{A}|\psi\rangle}{\langle\phi|\psi\rangle}. \quad (1.2)$$

Since the weak value is a kind of the expectation value, one would expect that it must take a real value and stay between the maximum and the minimum eigenvalues of the observable \hat{A} . However, the weak value can take an arbitrary complex number as we show later (see Section 3.3.1).

Since this cannot be possible to occur in classical probability theory, it must derive from the characteristics of quantum theory. This aspect has been studied intensively over the years from the interest in the foundations of quantum theory [9–13]. Despite that weak values have abnormal properties from the viewpoint of such ordinary statistical quantities, they are experimentally detectable quantities. In fact, it was confirmed experimentally by Ritchie *et al.* [14] for the first time in 1991 and since then there are numerous experiments verifying the physical existence of the weak value [15–17].

Meanwhile, it has also been attracted a lot of interest in its applicability in actual measurement. Due to the above-mentioned property, the weak value may take a value of the order larger than the typical eigenvalue of an observable \hat{A} characterizing the target physical phenomenon in the measurement. Accordingly, when there is any room for change that reflects the value of the observable \hat{A} measured experimentally, the change could be made large. The enhancement of the change is called *weak value amplification* (WVA).

In this respect, the detection experiment of the spin Hall effect of light by Hosten and Kwiat [18] in 2008 was a major breakthrough for the WVA to attract much attention for precision measurement. In this experiment, the shift due to the spin Hall effect of light, which was theoretically estimated to be on the order of 10 nanometers, was detected for the first time. Also, the authors have argued that the amplification effect is very important in its detection. Since the success of this experiment, the study on amplification using weak values has been persistently active both experimentally and theoretically.

On the experimental side, several experiments were conducted to study the property of the WVA by considering a number of different simple systems [19–22]. After these, many experiments followed to expand the range of the physical observable (*e.g.* temperature [23, 24], phase [25], velocity [26]) using optical systems. As more research has been done in recent years, more diversity is emerging. Specifically, we have seen examples such as a WVA experiment using the degree of freedom of an atom instead of an optical system [27] as well as an experiment for measuring variables of a chiral molecule [28].

On the theoretical side, there are three main areas of research aimed at exploiting the effects of the WVA. These are to identify the pros and cons of the WVA method, to propose variants that modify and improve the setup, and lastly to find new targets of application.

As to the first one, we have seen initially rather negative views on the WVA despite the experimental success [29–31]. Subsequent studies gave some positive results, but under special conditions [32–34]. In recent years, some advantages have been reported from a simple estimation theoretical point of view [35]. It has also been pointed out that it is important to consider environmental factors that cause errors in experiments [36, 37].

As for the second, many variants of the WVA have been proposed. Specifically, there are discussions on recovering statistics by collecting photons that failed in postselection [38], on assisting WVA using entanglement [39], on meter systems that handle time delays that occur when sending pulses [40, 41], and on changes in the passing probability of postselection without using a meter system [42, 43].

Thirdly, there are also studies looking for new applications [27, 28, 44–47]. Specifically, we have proposals for gravitational wave detection [44, 48], for precision measurement of the gravitational constant [46], for measuring the electric dipole moment of neutrons [47], and for simultaneous estimation of multiple variables [49].

Given this current status of research, we will address the following two issues in this dissertation. The first is to clarify the conditions or situations under which the WVA is effective, and the second is to expand the scope of application of the WVA method.

For the first issue, to clarify when the WVA is effective, we have adopted an earlier proposal by Lee and Tsutsui [36] which presented a scheme to discuss the uncertainty involved in the WVA systematically. They have proposed a method to evaluate the trade-off relationship between the WVA and the uncertainties caused by the WVA from the viewpoint of the relative uncertainty. The proposal was based on the viewpoint of measurement error theory. Based on an improved version of this method, we have analyzed the validity of the experimental results of the two preceding experiments, *i.e.*, the experiment of the detection of the spin Hall effect of light (SHEL) by Hosten and Kwiat [18] and the experiment of the ultrasensitive beam deflection (USBD) [19]. These experiments were conducted before the theoretical discussions by Lee and Tsutsui was presented.

As a result of this analysis, we confirmed when the weak value is made larger, both the statistical error and the approximation error become larger. When viewed in terms of relative uncertainties, defined as the ratio of these errors to the effective size of signals gained by the WVA, both the statistical and the approximation uncertainty tend to remain constant or increase as the weak value is made larger. These results agree with those obtained in [29–31] which adopt a different method of analysis from ours. However, we found that the actual errors in the experiments are much larger than the expected elements of the uncertainty caused by the WVA, which can be understood that the actual errors include some errors due to technical imperfections as pointed out in [36]. Comparing our analysis with the results of the experiments, we see that the WVA is, in fact, effective in reducing these errors caused by the imperfections.

For the second issue, we recall the fact that other than the optical systems, the WVA has been applied only to a few systems such as neutrons [50, 51] and atomic systems [27, 45]. The situation does not change significantly even if we include theoretical proposals [46, 52]. There are several reasons for this limited range of applications of the WVA. One of the concerns with the implementability of the postselection. Also, the coherence length of the system imposes another obstacle. As an attempt to overcome these problems, we consider the system of B mesons. The time evolution of the states of the B meson has been used in experiments of observing the violation of Bell's inequality [53], which is made possible on account of the coherence of the B meson system. Thus, if we can find a possible method of implementing the postselection, we may be able to measure the CP violating parameters based on the WVA more precisely than previously done.

In order to carry out the postselection, we consider the correspondence between the state before the decay and the state after the decay and show that under certain technical assumptions, specifically those mentioned in the final paragraph of Subsection 3.5.1, the operation of selecting a state for particles generated at the decay is equivalent to implementing the postselection for the B meson system. We also find that, if we choose the Hamiltonian operator for the observable \hat{A} , the imaginary part of the weak value is related to the lifetime and that this lifetime can be extended, in effect, by 2.6 times compared to the original lifetime of the B meson. In addition, we examine whether the WVA actually improves the measurement accuracy of the CP violating parameters. Our result indicates that the accuracy of the estimation of CP parameters may be improved by 20% by using the WVA method.

This dissertation is organized as follows. We first provide in Chapter 2 a review of the error theory related to precision measurement. Based on this, in Chapter 3, we review the discussion of indirect measurement in quantum measurement theory, the properties of the WVA method, and some of the premises of the WVA method, which is the subject of this dissertation. Then, in Chapter 4, in response to the proposal by Lee and Tsutsui [36], the WVA and measurement error in the experiment will be presented concretely based on our analysis [54]. Chapter 5 introduces the results of examining the measurement of CP violating parameters seen in the decay process of B mesons, which was carried out with the aim of applying such a WVA method to high-energy systems [55]. Chapter 6 is devoted to the conclusion of this dissertation. Detailed calculations and some related matters are covered in the appendices.

This dissertation is based on the published paper [54] and the submitted paper [55].

Chapter 2

Metrology and Precision Measurement

In order to implement precision measurement, we must assess the measurement error properly. Roughly speaking, precision measurement is measurement with maximum reduction of measurement error. Therefore, proposing a new precision measurement method amounts to offering an approach to reduce such measurement error maximally.

In this chapter, we discuss how error and uncertainty arise in measurement. Also, we briefly review some relating ideas of error and uncertainty in estimation theory. We use these ideas in our analysis later.

2.1 Measurement and theoretical uncertainties

We all know that the values obtained by experiments always contain some uncertainty. Here, we reconsider the notion of measurement uncertainty and its importance in experiments.

In the textbook of Rabinovich [56], the idea of the measurement uncertainty is related to the reproducibility of the experiment. Suppose that the results of experiments expected to be the same are actually different. In that case, it means that there is no reproducibility, but whether the results are the same should be argued only within the measurement uncertainty.

On the other hand, in classical physics, there exist a complete set of physical quantities which generally determine the state of the physical system. This determination of the state means that different physical quantities which do not belong to the set can be determined from the values of the physical quantities in the complete set.

In classical physics, we assume a particular value to an arbitrary physical quantity. We expect that measurement is an operation to find the value, and the theory is a way to obtain it by calculation. If this is the case, we may regard the value as the *true value* of the physical quantity. We, however, all know that this cannot be the case in reality.

In this dissertation, we refer to the term *error* as the difference between the true value and the value obtained by measurement or by theoretical calculation. If we can determine the difference in one way or another, we can obtain the true value by back-calculation, that is, by error correction. For the actual error in measurement/theory, such a determination is not possible in general, and all we can do is to make a loose estimate of the magnitude of the error. In this dissertation, we use the term ‘measurement uncertainty’ to refer to this uncertainty of the measured values. Also, we use the term ‘theoretical uncertainty’ to refer to this uncertainty of the theoretical calculation.

2.2 Random uncertainty and systematic uncertainty

One element of reproducibility is repeatability. It concerns with how much the measurement results coincide when the experimenter repeats the experiment with the same conditions. It is not uncommon for the actual measurement results to disagree, no matter how much effort the experimenters make to arrange the experimental conditions. The ‘error’ which appears in this way will be called *random uncertainty*.

In the actual experiments, there are some other factors which can affect the results of the measurement, such as the deviation of the origin of the pointer in the meter. When these factors are present, the measurement results will differ from experiments carried out in different environments, even if the true value is equal. The uncertainty which derives from these factors will be called *systematic uncertainty*.

In practice, the systematic uncertainty may be estimated by comparing the results with other results obtained in different experiments. This can be understood well if we recall the annual changes in the values published in the Review of Particle Physics by the Particle Data Group, which indicate that the results of precise measurements of various physical quantities can determine the systematic errors of the past experimental results. See Figure 2.1 for the annual transition of measurement results for several parameters appeared on page 19 of the 2020 edition [57].

In this way, the identification of systematic errors in precision measurements may take many years. In particular, in the case of high energy experiments, it usually takes several decades. Therefore, to reduce or to determine the systematic uncertainty in a relatively short period, it should be crucial to develop an entirely new method for the experiments.

2.3 Direct and indirect measurement in metrology

In general, the analysis of error depends on the processing method of the measurement results. The quantity we seek to obtain by measurement may be the quantity obtained directly by this measurement (‘direct measurement’) or the quantity estimated from other measurement results. (‘indirect measurement’) [56].

It should be noted that the meanings of ‘direct measurement’ and ‘indirect measurement’ in quantum measurement introduced in a later chapter are different from this. Only in this chapter, indirect measurement is defined as *a measurement in which the value of the measurand is calculated using measurements of other quantities related to the measurand by some known relation* [56]. On the other hand, direct measurement is a measurement in which the amount to be measured is obtained directly from the measuring instrument as a numerical value.

By definition, the indirect measurement can be represented by a function f which connects the measurement result (A_1, \dots, A_N) and quantity X we want to know as

$$X = f(A_1, \dots, A_N). \quad (2.1)$$

Depending on direct measurement or indirect measurement, handling the uncertainty differs. When we consider the direct measurement, the measurement uncertainty is equal to the uncertainty of the quantity that we want to know. On the other hand, in indirect measurement, it is necessary to consider error propagation

$$\delta_i X = \frac{\partial f}{\partial A_i} \delta A_i, \quad (2.2)$$

where $\delta_i X$ is the error of X caused by the error δA_i of A_i .

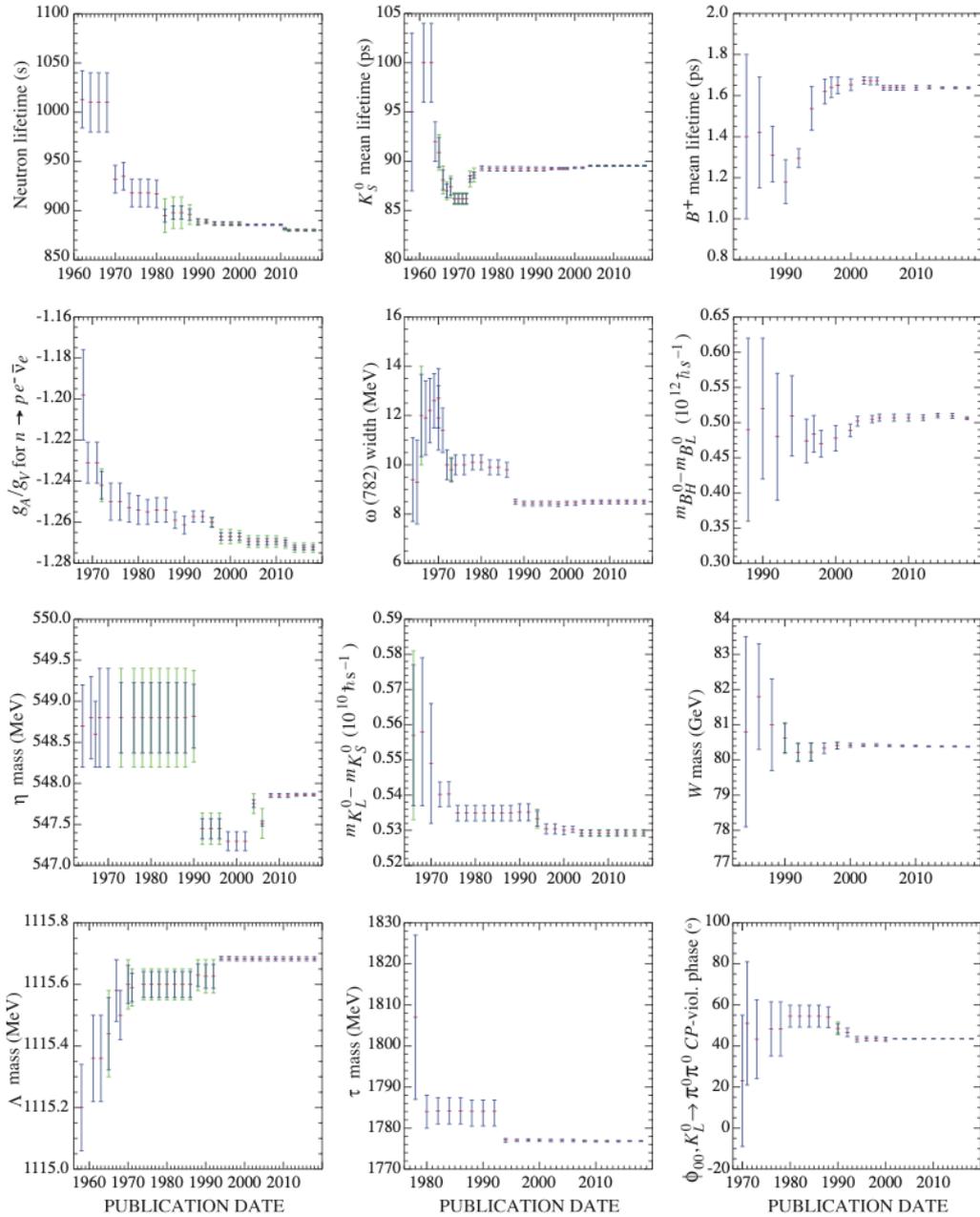


Figure 1: A historical perspective of values of a few particle properties tabulated in this Review as a function of date of publication of the Review. A full error bar indicates the quoted error; a thick-lined portion indicates the same but without the “scale factor.”

Figure 2.1: Fig. 1 on P.19 of [57]. The measurement results of various parameters have been changing with the improvement of experiments. Although the accuracy of the results has improved over the years, systematic errors are often found by the results of later years.

In some cases, the relationship (2.1) may be available only stochastically. In that case, we can use an expected relationship motivated by stochastic behaviors. The prescriptions applicable for this are combined into a method called *parameter estimation*.

2.4 Parameter estimation and its uncertainty

There are many discussions related to probability in measurement, not limited to discussions involving quantum theory. Besides, the probability distributions that appear in those discussions may include parameter(s) that cannot be determined by calculation alone. We often need to estimate such stochastic parameter(s) using the results of experiments.

In this case, since predictions based on the probability theory are meaningful after infinitely many trials. Accordingly, it is inevitable to have some uncertainty when we use the predictions to establish a relationship between the results obtained by a finite number of measurements and the parameter in a probability distribution. Two well-known methods for estimating the parameter(s) with uncertainty are the method of *least squares* and *maximum likelihood estimation*. Since the latter one is used in Chapter 5 of this dissertation, we briefly review it below.

2.4.1 The maximum likelihood estimation method

The maximum likelihood estimation (MLE) method is a well-used method to estimate parameters included in the theoretically given probability distribution.

Estimation of the parameters included in the probability distribution

Suppose that, given a probability distribution $P(x|\theta)$ where θ is an unknown constant and x is a random variable, we wish to estimate this θ experimentally.

By repeating the experiment N times, we can obtain corresponding measurement results for the random variable x , which we describe $\{x_1, \dots, x_N\}$.

Then, we define a function of θ as

$$L(\theta) := \prod_{i=1}^N P(x_i|\theta). \quad (2.3)$$

This is called the likelihood function. The direct meaning of this function is the probability of obtaining a combination of values (x_1, \dots, x_N) by performing N independent trials that can be described by a probability distribution $p(x|\theta)$ characterized by a fixed θ . In the MLE method, a value

$$\theta_{\text{est}} = \operatorname{argmax} L(\theta), \quad (2.4)$$

which is the value of θ for which $L(\theta)$ takes the maximal value, is assigned as an estimated value for the constant θ .

Uncertainty of the estimated value (Neyman method)

Various factors cause the uncertainty for the estimated value θ . We classify this uncertainty into two parts. One is the statistical uncertainty, which derives from the insufficient convergence of the distribution due to the small number of trials N . The other is the systematic uncertainty, which is the measurand change caused by some effects characterized

by different parameters called nuisance parameters. When these nuisance parameters and θ are correlated, the systematic error of θ arises from the non-zero nuisance parameters. We can modify our theoretical prediction for the measurement results if we know the probability distribution for the nuisance parameters.

We now evaluate the statistical effects that remain after taking into account the modification of such probability distributions. Neyman proposed the following method to assess the uncertainty of the parameter θ [58].

Given a probability distribution $P(x|\theta)$, let us determine $x_1(\theta, \alpha)$ and $x_2(\theta, \alpha)$ such that

$$P(x_1 < x < x_2|\theta) := \int_{x_1}^{x_2} P(x|\theta)dx = 1 - \alpha. \quad (2.5)$$

Since the condition of coverage in (2.5) does not determine $x_1(\theta, \alpha)$ and $x_2(\theta, \alpha)$ uniquely, some additional condition is needed. One possibility is to choose central intervals such that the probabilities of finding x below x_1 and above x_2 are each $\alpha/2$. To avoid complexity, we assume that the functions $x_1(\theta, \alpha)$ and $x_2(\theta, \alpha)$ are both monotonous and smooth with respect to θ for any $\alpha \in (0, 1]$.

Here, we consider the inverse functions for $x_1(\theta, \alpha)$ and $x_2(\theta, \alpha)$, and write them as $\theta_1(x, \alpha)$ and $\theta_2(x, \alpha)$, respectively. This means that, for any possible x ,

$$x_1(\theta_1(x, \alpha), \alpha) = x, \quad (2.6)$$

and

$$x_2(\theta_2(x, \alpha), \alpha) = x. \quad (2.7)$$

Using the properties of the inverse function and the fact that $x_1(\theta, \alpha) < x_2(\theta, \alpha)$, we have $\theta_1(x, \alpha) > \theta_2(x, \alpha)$ in the case that both $x_1(\theta, \alpha)$ and $x_2(\theta, \alpha)$ are monotonically increasing. Also, $\theta_1(x, \alpha) < \theta_2(x, \alpha)$ in the case that both $x_1(\theta, \alpha)$ and $x_2(\theta, \alpha)$ are monotonically decreasing. To see this, we first consider the equation

$$x_1(\theta_1, \alpha) = x_2(\theta_2, \alpha) = x, \quad (2.8)$$

and define the difference

$$\delta\theta = \theta_2 - \theta_1. \quad (2.9)$$

When the difference (2.9) is applied to (2.8), we can obtain

$$x_1(\theta_1, \alpha) = x_2(\theta_1 + \delta\theta, \alpha) = x_2(\theta_1, \alpha) + \delta x_2, \quad (2.10)$$

where δx_2 is

$$\delta x_2 := x_2(\theta_2, \alpha) - x_2(\theta_1, \alpha). \quad (2.11)$$

Due to $x_1(\theta, \alpha) < x_2(\theta, \alpha)$, we find

$$\delta x_2 < 0. \quad (2.12)$$

If $x_2(\theta, \alpha)$ is a monotonically increasing function of θ , then $\theta_1 > \theta_2$. Also, if $x_2(\theta, \alpha)$ is a monotonically decreasing function of θ , then $\theta_1 < \theta_2$ as we stated above.

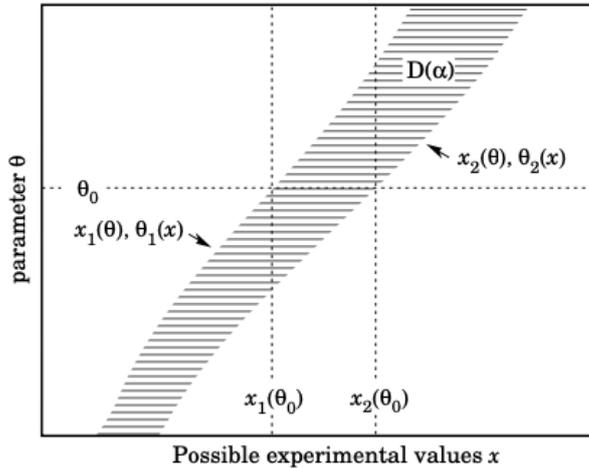


Figure 2.2: Figure 40.3 from [57]. For each θ , the width of the distribution of the measured quantity x is given as $D(\alpha)$. This $D(\alpha)$ is in the form of a belt, which is called the ‘confidence belt’ in [57]. The treatment that the width of this confidence belt at x is considered as a confidence interval for θ when a certain measurement result x is obtained.

Assuming that x is monotonically increasing with respect to θ , if we consider a function $\tilde{P}(\theta_2(x, \alpha) < \theta < \theta_1(x, \alpha))$ satisfying

$$1 - \alpha = P(x_1 < x < x_2 | \theta) = \tilde{P}(\theta_2(x, \alpha) < \theta < \theta_1(x, \alpha)), \quad (2.13)$$

the corresponding interval $[\theta_1, \theta_2]$ may be used as a confidence interval. On the other hand, if x_1 and x_2 are monotonically decreasing with respect to θ , we have the condition (2.13) with θ_1 and θ_2 interchanged. This prescription is illustrated in Figure 2.2.

We note that the present prescription of estimation requires the assumption that the random variable x obtained by the measurement is a ‘true value.’ In other words, it does not take into account the uncertainty of the measurement of the variable x .

2.5 Summary

It is known that the numerical values obtained in the experiments have uncertainty as a quantity that characterizes the reproducibility of the results. It is crucial to remove such uncertainty in a reproducible manner, which is the essence of precision measurement.

The processing of such uncertainties also depends on the relationship between the estimated value and the measured value. A particular method called parameter estimation is required when probability theory is involved. Due to the nature of the systematic error present in the experimental results, it is necessary to construct a set of different environment and system for specifying the systematic uncertainty.

Chapter 3

Weak Value Amplification

In this chapter, we briefly review the weak value amplification (WVA) and its related concepts.

Section 3.1 is devoted to explaining the model of indirect measurement, in particular, the von Neumann measurement model. The original WVA method is realized by supplementing the postselection to the von Neumann model. This measurement is called ‘Aharonov’s weak measurement’ or simply ‘weak measurement’. Below we will explain the Aharonov’s weak measurement in Section 3.2 and its amplification effect in Section 3.3. Although the WVA method has already been implemented in various experiments, it also has been pointed out that there are some controversial points with the method. These points make the proposed merit of the WVA unclear and give us some cautions for using the WVA. These points will be mentioned in Section 3.4.

Finally, in Section 3.5, we provide some technical remarks relevant to actual experiments. These will be important for us to analyze the actual systems and explore a new application method.

3.1 Indirect measurement in quantum measurement

In quantum mechanics, the projection postulate is an essential ingredient in describing quantum measurement. The postulate insists that, when we obtain a as a result of measuring on a physical quantity A represented by the quantum observable \hat{A} , the value a must be one of the eigenvalues of the operator \hat{A} and that the state of the system becomes the eigenstate $|a\rangle$ corresponding to the eigenvalue a .

The postulate leads us to conclude that all the same measurement results are identical when we repeat the measurement in exactly the same way. However, the actual experiments may produce different values within the range of measurement uncertainty. Therefore, we find that the projection postulate cannot strictly describe the actual experiment. The idea of indirect measurement appeared as one of the means to cope with such an actual measurement process mathematically.

In actual measurements we often proceed as follows: First, we prepare a measurable quantity X that may not necessarily be related to the physical quantity A we want to measure. Next, we realize a correlation between the quantities A and X through some interaction. Finally, we measure the quantity X . For instance, in the case of the Stern-Gerlach experiment, observers intend to measure the spin of the particles, but what they directly measure is the directions in which the particles are bent in the apparatus or the position where the particles reach. In either case, what we actually measure is not what we want to measure. Such a measurement is called *indirect measurement*.

To be more specific, let \mathcal{H} be the Hilbert space in which the observable \hat{A} is represented. We call the system of \mathcal{H} the ‘target system’. Similarly, let \mathcal{K} be the Hilbert space in which the observable \hat{X} is represented. We call the system of \mathcal{K} the ‘meter system’. The correlation we want to realize will then be provided by the interaction between the target system and the meter system. In short, the indirect measurement is characterized by the target system, the meter system, and the interaction.

3.1.1 The von Neumann measurement model

The von Neumann measurement model is well established as a model of indirect measurement. In this model, the simple one-dimensional quantum system is considered for the meter system. This system is given by the space of square-integrable functions $\mathcal{K} = L^2(\mathbb{R})$, where a pair of observables \hat{X} and \hat{P} satisfying the canonical commutation relation (CCR),

$$[\hat{X}, \hat{P}] = i, \quad (3.1)$$

are equipped. The interaction of the von Neumann measurement model is furnished by the unitary operator acting in the total system $\mathcal{H} \otimes \mathcal{K}$,

$$\hat{U}(\theta) := e^{-i\theta\hat{A}\otimes\hat{Y}}, \quad (3.2)$$

where θ is a real parameter representing the coupling of the target system and the meter system, and \hat{A} is an observable of the target system. The von Neumann measurement describes an indirect measurement on the quantity corresponding to this observable \hat{A} . For our convenience, we introduce the notation \hat{Y} as an operator of the meter system to represent either $\hat{Y} = \hat{X}$ or $\hat{Y} = \hat{P}$.

To see that these setups can describe a measurement, let $|\chi\rangle$ be the state of the meter system prepared initially along with the state of the target system $|\psi\rangle$. Because the state $|\chi\rangle$ is an element of $L^2(\mathcal{R})$, we can express it as

$$|\chi\rangle = \int dx \chi(x)|x\rangle, \quad (3.3)$$

where $\chi(x)$ is often called a wave function and $|x\rangle$ is a vector satisfying $\langle x'|x\rangle = \delta(x' - x)$. Then, the initial state of the total system is the product state

$$|\Psi\rangle = |\psi\rangle|\chi\rangle. \quad (3.4)$$

Under the measurement interaction (3.2), the state of the total system (3.4) undergoes the change,

$$\begin{aligned} |\Psi\rangle &\rightarrow |\Phi\rangle := \hat{U}(\theta) |\psi\rangle|\chi\rangle, \\ &= e^{-i\theta\hat{A}\otimes\hat{Y}} |\psi\rangle|\chi\rangle. \end{aligned} \quad (3.5)$$

The state $|\psi\rangle$ can be expanded using a series of complex coefficients $\{c_i\}$ as

$$|\psi\rangle = \sum_i c_i |a_i\rangle, \quad (3.6)$$

where $|a_i\rangle$ is an eigenstate of the observable \hat{A} satisfying $\hat{A}|a_i\rangle = a_i|a_i\rangle$. Applying (3.6) to (3.5), we find

$$\begin{aligned} |\Phi\rangle &:= \sum_i e^{-i\theta\hat{A}\otimes\hat{Y}} (c_i |a_i\rangle) |\chi\rangle, \\ &= \sum_i c_i |a_i\rangle |\chi_i\rangle, \end{aligned} \quad (3.7)$$

where $|\chi_i\rangle$ is defined by

$$|\chi_i\rangle := e^{-i\theta a_i \hat{Y}} |\chi\rangle. \quad (3.8)$$

When the operator \hat{Y} is the momentum operator \hat{P} , the state (3.8) can be understood as a state obtained by translating the initial state $|\chi\rangle$ in the + direction by θa_i . In fact, for any state $|\chi\rangle$ of the meter system with the corresponding wave function $\langle x|\chi\rangle = \chi(x)$ and an arbitrary real number θ , we obtain

$$e^{-i\theta \hat{P}} |\chi\rangle = e^{-i\theta \hat{P}} \int_{-\infty}^{\infty} dx \chi(x) |x\rangle = \int_{-\infty}^{\infty} dx \chi(x - \theta) |x\rangle, \quad (3.9)$$

(see Appendix B.1 for details) on account of the fact that the momentum operator \hat{P} is the generator of the translation.

It shows that the state of the meter system varies according to each eigenstate of the target observable \hat{A} . This entanglement enables us to perform a measurement on the target system through the measurement on the meter system.

3.1.2 Weakness of the measurement

When the states $|\chi_i\rangle$ defined by (3.5) satisfy

$$\langle \chi_i | \chi_j \rangle = \begin{cases} 1 & (i = j), \\ 0 & (i \neq j), \end{cases} \quad (3.10)$$

we can determine the state of the target system by measuring in which $|\chi_i\rangle$ the state of the meter system is. In this case, this indirect measurement reduces to the standard projection measurement. Actually, when the result of the measurement on the meter system is a_k and its corresponding state of the meter system is $|\chi_k\rangle$, the combined state after the measurement becomes $|a_k\rangle|\chi_k\rangle$.

However, the condition (3.10) is not always satisfied. In fact, since the inner product is

$$\begin{aligned} \langle \chi_i | \chi_j \rangle &= \langle \chi | e^{i\theta a_i \hat{Y}} e^{-i\theta a_j \hat{Y}} | \chi \rangle \\ &= \langle \chi | e^{i\theta(a_i - a_j) \hat{Y}} | \chi \rangle, \end{aligned} \quad (3.11)$$

we find that the inner product $\langle \chi_i | \chi_j \rangle$ depends on the initial state of the meter system $|\chi\rangle$, the parameter θ , the set of eigenvalues of \hat{A} , and the observable of the meter system \hat{Y} .

When there is no degeneracy in the target observable \hat{A} , the operator $e^{i\theta(a_i - a_j) \hat{Y}}$ can be made significantly different from the identity operator by making θ large enough. Then, the condition (3.10) may hold approximately. On the other hand, $e^{i\theta(a_i - a_j) \hat{Y}}$ approaches the identity when θ is sufficiently small. When θ is precisely zero, there is no interaction between the two systems, and accordingly no change occurs in the meter system. In this case, the indirect measurement model does not work.

The parameter θ characterizes that the strength of the interaction, and we call the limit where θ is close to zero ‘weak limit’.

3.1.3 Measurement on the meter system

Suppose that the operator \hat{X} corresponding to the pointer variable X is the actual position operator. Suppose also that the observable \hat{Y} in (3.2) is the momentum operator \hat{P} . Then, the interaction operator (3.2) acts as the translation operator for the pointer variable X of the meter system by the amount of the value of the target observable \hat{A} of the target system.

Recall that a unitary operator $e^{i\delta\hat{P}}$ is an operator for translation, that is,

$$e^{i\delta\hat{P}}\hat{X}e^{-i\delta\hat{P}} = \hat{X} + \delta, \quad (3.12)$$

due to the Baker-Campbell-Hausdorff formula,

$$e^{\hat{P}}\hat{X}e^{-\hat{P}} = e^{ad_{\hat{P}}}\hat{X} := \hat{X} + [\hat{P}, \hat{X}] + \frac{1}{2!}[\hat{P}, [\hat{P}, \hat{X}]] + \dots, \quad (3.13)$$

with (3.1). When we perform a measurement on the meter system of the state (3.7) for the pointer variable, the result is given by the probability distribution,

$$P(x) = \langle \Phi | (\hat{1} \otimes |x\rangle\langle x|) | \Phi \rangle = \sum_i |c_i|^2 |\langle x | \chi_i \rangle|^2 = \sum_i |c_i|^2 P_\chi(x - \theta a_i), \quad (3.14)$$

where

$$P_\chi(x) = |\langle x | \chi \rangle|^2, \quad (3.15)$$

is the probability distribution of the result when the pointer variable is measured for the state before the interaction. This $P(x)$ in (3.14) can be regarded as a distribution formed by probabilistic mixing of the translated distributions.

The expectation value of the pointer variable after the interaction then reads

$$\langle \Phi | \hat{1} \otimes \hat{X} | \Phi \rangle = \langle \Psi | e^{i\theta\hat{A} \otimes \hat{Y}} (\hat{1} \otimes \hat{X}) e^{-i\theta\hat{A} \otimes \hat{Y}} | \Psi \rangle = \langle \Psi | \hat{1} \otimes \hat{X} | \Psi \rangle + \theta \langle \psi | \hat{A} | \psi \rangle. \quad (3.16)$$

The difference between the expectation value of the pointer before and after the interaction is then found as

$$\Delta(\theta) := \langle \Phi | \hat{1} \otimes \hat{X} | \Phi \rangle - \langle \Psi | \hat{1} \otimes \hat{X} | \Psi \rangle = \theta \langle \psi | \hat{A} | \psi \rangle. \quad (3.17)$$

We note that this shift (3.17) is small when the parameter θ is small. We also note that changing the state $|\psi\rangle$ does not significantly change the shift $\Delta(\theta)$ (see Fig. 3.1). Consequently, we now can retrieve the expectation value of the target observable \hat{A} directly from the shift of the meter as $\Delta(\theta)/\theta = \langle \psi | \hat{A} | \psi \rangle$.

These results form the basis for our comparison with the weak measurement which we now describe below.

3.2 Aharonov's weak measurement

After the interaction (3.7), the state is entangled, in which case we specify the state of the target system by measuring the meter system, and vice versa. The basic idea of Aharonov's weak measurement is that, in addition to the conventional process called *pre-selection* in which we prepare the initial state $|\psi\rangle$ for the target system, we also implement the process called *postselection* before we observe the meter system.

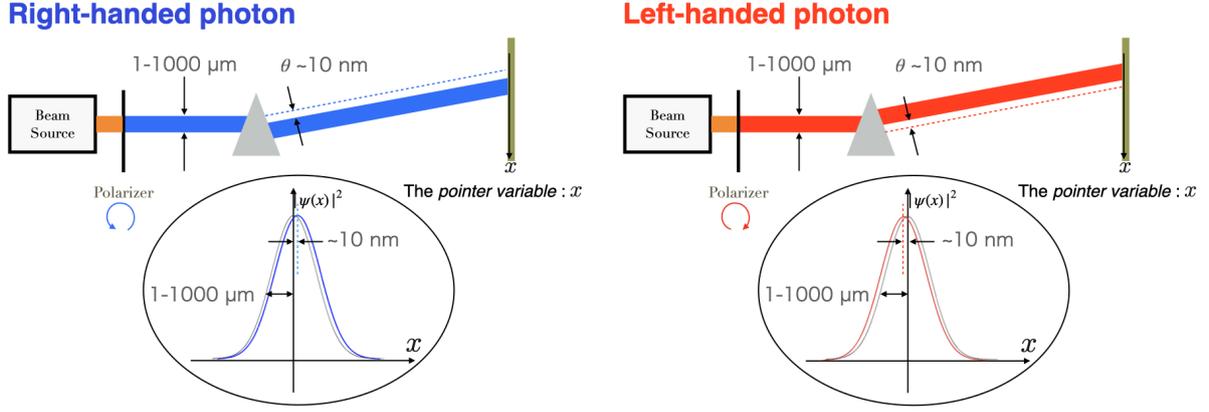


Figure 3.1: These figures describe the experiment for detecting the spin Hall effect described later as an indirect measurement. In this case, a space of photon spin wave functions is assigned to \mathcal{H} , and space of wave functions for the photon position x on the rightmost screen is assigned to \mathcal{K} . The distribution shown at the bottom of each figure on both the left and right sides is a probability distribution showing the probability density where photons are detected. The figure on the left shows that we use only right-handed photons. The distribution is shifted in the $+$ direction. Also, the figure on the right shows that we use only left-handed photons. The distribution is shifted in the $-$ direction.

Once we choose a state $|\phi\rangle$ of the target system for the postselected state, we examine if the state of the target system is actually $|\phi\rangle$ or not after the interaction (3.2). Measuring the pointer variable is carried out only when the result of the examination is affirmative, in which case the (normalized) state of the meter system becomes

$$|\xi\rangle = \frac{\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle}{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|}, \quad (3.18)$$

$$= \sum_i \frac{c_i \langle\phi|a_i\rangle}{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|} e^{-i\theta a_i \hat{Y}} |\chi\rangle. \quad (3.19)$$

The difference between the initial state $|\chi\rangle$ of the meter system and the state $|\xi\rangle$ given by (3.18) can be characterized by a value called the weak value as shown below.

To see this, we first expand the operator $\hat{U}(\theta)$ in (3.18), given by (3.2), up to the first order of the parameter θ . Then, the state (3.18) becomes

$$\begin{aligned} |\xi\rangle &\propto \langle\phi|(\hat{1} - i\theta\hat{A} \otimes \hat{Y})|\psi\rangle|\chi\rangle \\ &= \langle\phi|\psi\rangle(\hat{1} - i\theta A_w \hat{Y})|\chi\rangle \\ &\simeq \langle\phi|\psi\rangle e^{-i\theta A_w \hat{Y}} |\chi\rangle, \end{aligned} \quad (3.20)$$

where A_w is the weak value defined by

$$A_w := \frac{\langle\phi|\hat{A}|\psi\rangle}{\langle\phi|\psi\rangle}. \quad (3.21)$$

When the operator \hat{Y} is the momentum \hat{P} , the state (3.20) is equivalent to translating the initial state $|\chi\rangle$ by the distance θA_w due to the result (3.9).

The absolute square of the denominator of (3.18),

$$q := |\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|^2, \quad (3.22)$$

gives the probability that a specific state $|\phi\rangle$ can be obtained by performing the postselection on the target system \mathcal{H} according to the Born rule. We call this q in (3.22) the *postselection rate*.

We can obtain a distribution of the results of measurements of the pointer variable on the state (3.19) with the transition amplitude

$$\langle x|\xi\rangle = \frac{\langle\phi|\langle x|\hat{U}(\theta)|\psi\rangle|\chi\rangle}{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|}, \quad (3.23)$$

$$= \sum_i \frac{c_i \langle\phi|a_i\rangle}{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|} \langle x|e^{-i\theta a_i \hat{Y}}|\chi\rangle. \quad (3.24)$$

In the case $\hat{Y} = \hat{P}$, (3.24) reads

$$\langle x|\xi\rangle = \sum_i \frac{c_i \langle\phi|a_i\rangle}{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|} \chi(x - \theta a_i). \quad (3.25)$$

This means that the translated wave functions are superposed by coefficients determined by the postselection. This is different from the mixture of probability distributions seen in the indirect measurement without postselection (3.14). Also, it is expected that the amplitude can be canceled depending on the coefficients.

A peculiar result caused by the postselection can be seen not only in the distribution shape but also in the average value of the measurement results of the pointer variable. To see this, we calculate the expectation value

$$\langle\xi|\hat{X}|\xi\rangle = \langle\chi|\frac{\langle\psi|\hat{U}(\theta)|\phi\rangle\hat{X}\langle\phi|\hat{U}(\theta)|\psi\rangle}{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|^2}|\chi\rangle. \quad (3.26)$$

In this case, instead of the relation (3.12), we use the Taylor expansion of the unitary operator (3.2) up to the linear order in θ to obtain

$$\begin{aligned} \langle\xi|\hat{X}|\xi\rangle &\simeq \langle\chi|\frac{\langle\psi|(\hat{1} + i\theta\hat{A} \otimes \hat{Y})|\phi\rangle\hat{X}\langle\phi|(\hat{1} - i\theta\hat{A} \otimes \hat{Y})|\psi\rangle}{|\langle\phi|(\hat{1} - i\theta\hat{A} \otimes \hat{Y})|\psi\rangle|\chi\rangle|^2}|\chi\rangle \\ &\simeq \langle\chi|\frac{(\hat{X} + i\theta A_w^* \hat{Y} \hat{X} - i\theta A_w \hat{X} \hat{Y})}{\langle\chi|(1 + i\theta A_w^* \hat{Y} - i\theta A_w \hat{Y})|\chi\rangle}|\chi\rangle \\ &\simeq \left(\langle\chi|\hat{X}|\chi\rangle + 2\theta \text{Im}[A_w \langle\chi|\hat{X}\hat{Y}|\chi\rangle]\right) \cdot \left(1 - 2\theta \text{Im}[A_w \langle\chi|\hat{Y}|\chi\rangle]\right) \\ &= \langle\chi|\hat{X}|\chi\rangle + \theta \cdot 2\text{Im}[A_w C(\hat{X}, \hat{Y})], \end{aligned} \quad (3.27)$$

where $C(\hat{X}, \hat{Y})$ is a quantum version of covariance defined by

$$C(\hat{X}, \hat{Y}) := \langle\chi|\left(\hat{X} - \langle\chi|\hat{X}|\chi\rangle\right)\left(\hat{Y} - \langle\chi|\hat{Y}|\chi\rangle\right)|\chi\rangle, \quad (3.28)$$

$$= \langle\chi|\hat{X}\hat{Y}|\chi\rangle - \langle\chi|\hat{X}|\chi\rangle\langle\chi|\hat{Y}|\chi\rangle. \quad (3.29)$$

From this, we learn that the shift of the pointer variable \hat{X} in the weak measurement is indeed given by the formula:

$$\begin{aligned} \Delta_w(\theta) &:= \langle\xi|\hat{X}|\xi\rangle - \langle\chi|\hat{X}|\chi\rangle, \\ &= \theta \cdot 2\text{Im}\left[A_w C(\hat{X}, \hat{Y})\right] + O(\theta^2). \end{aligned} \quad (3.30)$$

Using the identity

$$\hat{X}\hat{Y} = \frac{1}{2} \left(\{\hat{X}, \hat{Y}\} + [\hat{X}, \hat{Y}] \right), \quad (3.31)$$

and the hermiticity of the anti-commutation relation together with anti-hermiticity of the commutation relation, we find

$$\Delta_w(\theta) = \theta \cdot \left\{ \text{Re}[A_w] \frac{1}{i} \langle \chi | [\hat{X}, \hat{Y}] | \chi \rangle + \text{Im}[A_w] \left(\langle \chi | \{\hat{X}, \hat{Y}\} | \chi \rangle - 2 \langle \chi | \hat{X} | \chi \rangle \langle \chi | \hat{Y} | \chi \rangle \right) \right\} \quad (3.32)$$

in the weak limit ($\theta \ll 1$). In particular, when $\hat{Y} = \hat{P}$, the shift (3.32) becomes

$$\Delta_w(\theta) = \theta \cdot \left\{ \text{Re}[A_w] + \text{Im}[A_w] \left(\langle \chi | \{\hat{X}, \hat{P}\} | \chi \rangle - 2 \langle \chi | \hat{X} | \chi \rangle \langle \chi | \hat{P} | \chi \rangle \right) \right\}. \quad (3.33)$$

The term proportional to $\text{Re}[A_w]$ is independent of the state of the meter system. When we can ignore the term proportional to $\text{Im}[A_w]$, we find that $\text{Re}[A_w]$ corresponds to $\langle \psi | \hat{A} | \psi \rangle$ in (3.17). This result supports the idea that the weak value A_w can be considered as the expectation value of the observable \hat{A} when the preselected state $|\psi\rangle$ and the postselected state $|\phi\rangle$ are given [8, 59, 60]. In this way, the real part of the weak value $\text{Re}[A_w]$ has the property of a quantity corresponding to an observable \hat{A} when preselection and postselection are both performed.

For the sake of convenience later, we shall provide the result for the case $\hat{X} = \hat{Y}$, in which case we find

$$\begin{aligned} \Delta_w(\theta) &= 2\theta \text{Im}[A_w] \left(\langle \chi | \hat{X}^2 | \chi \rangle - \langle \chi | \hat{X} | \chi \rangle^2 \right) \\ &= 2\theta \text{Im}[A_w] \text{Var}(\hat{X}). \end{aligned} \quad (3.34)$$

The ‘weak’ in the name comes from the weak limit of the measurement. However, we stress that the weak value itself does not include the weakness of measurement in the definition. Also, it is known that under certain conditions the weak value appears regardless of the strength of measurement, for example, when we examine a measurement model other than the von Neumann measurement model [51].

3.3 Weak value amplification (WVA)

So far, we have seen the effect of the postselection in the indirect measurement in view of the distribution and the mean value of the measurement result of the pointer variable for the meter system. Also, we have confirmed that the shift of the expectation value of the pointer variable includes a weak value when postselection is performed.

Although the weak value is a kind of conditional average, it can take anomalous values as we show in Subsection 3.3.1. Due to this reason, even if θ is very small (that is, the ‘weak limit’ mentioned in the above discussion), a not-so-small change of the pointer variable will appear depending on the value of the weak value A_w . This feature forms the basis of weak value amplification and is well utilized in many experiments.

3.3.1 On the possible weak value

Although the weak value A_w is a kind of the expectation value of \hat{A} , the weak value can take an arbitrary complex number by selecting the preselected and postselected states appropriately, regardless of the range of the eigenvalues. First, we prove this.

Let \hat{A} be a target observable on \mathcal{H} with $\dim \mathcal{H} \geq 2$. One then may choose two states $|\psi\rangle, |\psi^\perp\rangle \in \mathcal{H}$ fulfilling

$$\langle\psi|\psi^\perp\rangle = 0, \quad \langle\psi|\hat{A}|\psi^\perp\rangle \neq 0. \quad (3.35)$$

Given any $z \in \mathbb{C}$, if we choose

$$|\phi(z)\rangle := \frac{|\psi\rangle + z^* \cdot |\psi^\perp\rangle}{\sqrt{\|\psi\|^2 + |z|^2\|\psi^\perp\|^2}}, \quad (3.36)$$

for our postselected state $|\phi\rangle = |\phi(z)\rangle$, we find that the weak value becomes

$$A_w(z) = \frac{\langle\phi(z)|\hat{A}|\psi\rangle}{\langle\phi(z)|\psi\rangle} = \langle\psi|\hat{A}|\psi\rangle + z \cdot \langle\psi^\perp|\hat{A}|\psi\rangle. \quad (3.37)$$

Since we can choose z arbitrarily, the weak value may take an arbitrary value, implying that the pointer variable's shift can be amplified as much as we wish. To take advantage of the property, we are required to prepare appropriate superposition states by the preselection and postselection.

This kind of amplification effect is introduced in the paper [8] in which the weak value was first proposed as

Another striking aspect of this experiment becomes evident when we consider it as a device for measuring a small gradient of the magnetic field $\partial B_z/\partial z$. Our choosing α close to π yields a tremendous amplification.

To mention a few characteristics of the weak value, we first note that when either the preselected state $|\psi\rangle$ or the postselected state $|\phi\rangle$ is an eigenstate of the observable \hat{A} , the weak value A_w becomes its corresponding eigenvalue. Second, if both the preselected state and the postselected state are not eigenstates of \hat{A} , these states must not be orthogonal to each other to define the weak value.

3.3.2 ‘Tail’ can be more affected from the interaction

According to the discussion given so far, the combination of the preselection and postselection of the target system can lead to a particular amplification effect due to the appearance of weak values in the shift. We now show another point of view for the WVA.

In calculating the transition amplitude, we use an analogous process in the postselection of the target system and in the measurement of the pointer variable in the meter system, and it does not matter which one should be performed first. However, in the above discussion, we adopted the order of first selecting the target system and then discussing the corresponding results of the measurement of the meter system. In this section, we shall do the opposite to gain a new point of view for the amplification mechanism.

When we obtain a result x in the measurement of the pointer variable on the state before the postselection (3.7), the state of the target system becomes

$$\langle x|\Phi\rangle = \sum_i c_i \langle x|\chi_i\rangle |a_i\rangle, \quad (3.38)$$

up to normalization. We then observe that the measurement of the meter system on the entangled total system transforms the prepared initial state into the state (3.38). To see

this, for definiteness, we choose $\hat{Y} = \hat{P}$ to find

$$\begin{aligned}
|\tilde{\psi}\rangle &:= \langle x|\Phi\rangle = \sum_i c_i \chi(x - \theta a_i) |a_i\rangle \\
&= \sum_i c_i (\chi(x) - \theta a_i \chi'(x) + O(\theta^2)) |a_i\rangle \\
&= \chi(x) \left(\sum_i c_i |a_i\rangle - \theta \frac{\chi'(x)}{\chi(x)} \sum_i c_i a_i |a_i\rangle \right) + O(\theta^2) \\
&= \chi(x) \left(|\psi\rangle - \theta \frac{\chi'(x)}{\chi(x)} \hat{A} |\psi\rangle \right) + O(\theta^2).
\end{aligned} \tag{3.39}$$

The transition amplitude between this state and the postselected state $|\phi\rangle$ reads

$$\langle \phi|\tilde{\psi}\rangle = \chi(x) \left(\langle \phi|\psi\rangle - \theta \frac{\chi'(x)}{\chi(x)} \langle \phi|\hat{A}|\psi\rangle \right) + O(\theta^2). \tag{3.40}$$

We find that the first order term of θ in (3.40) includes the ratio of $\chi(x)$ and its derivative $\chi'(x)$. This behavior of the ratio depends on the function $\chi(x)$, but in the case where $\chi(x)$ is a Gaussian or a function obtained by multiplying the Gaussian by a polynomial, it can be seen that outliers take larger values. Indeed, if

$$\chi(x) = p(x)e^{-kx^2}, \tag{3.41}$$

where $p(x)$ is a polynomial function of x and k is a constant, the ratio is given by

$$\frac{\chi'(x)}{\chi(x)} = \frac{(p'(x) - 2kxp(x))e^{-kx^2}}{p(x)e^{-kx^2}} = \frac{p'(x)}{p(x)} - 2kx. \tag{3.42}$$

It can be seen that the larger the absolute value of x is, the larger the ratio becomes. The WVA is a method to obtain outliers selectively in the distribution and to remove other parts line in the center.

Figure 3.2 illustrates the case that the postselection is added compared to the case that the postselection is absent depicted in Figure 3.1.

3.3.3 Other schemes for deriving the weak value

In recent years, different methods for detecting a weak value have been proposed without using the above method [42,43]. Since they share an intention to obtain some amplification effect by making the weak value larger, it is also called WVA.

To introduce them, we first introduce the following small linear transformation $\hat{N}(\theta)$, which is a normal operator parametrized by a tiny parameter θ . In this subsection we choose θ to be a complex value. We assume that $\hat{N}(\theta)$ satisfies $\hat{N}(0) = \hat{1}$ and can be expanded for θ as

$$\hat{N}(\theta) = \hat{1} + \theta \hat{C} + O(\theta^2), \tag{3.43}$$

where \hat{C} is the derivative of $\hat{N}(\theta)$ at $\theta = 0$, and we call \hat{C} the *generator* of $\hat{N}(\theta)$. We consider the change in the postselection rate, which is the passing probability of the postselection given by (3.22) in the ordinary WVA case, due to this transformation $\hat{N}(\theta)$,

$$|\langle \phi|\hat{N}(\theta)|\psi\rangle|^2 = |\langle \phi|\psi\rangle|^2 (1 + 2\text{Re}[\theta C_w]) + O(\theta^2). \tag{3.44}$$

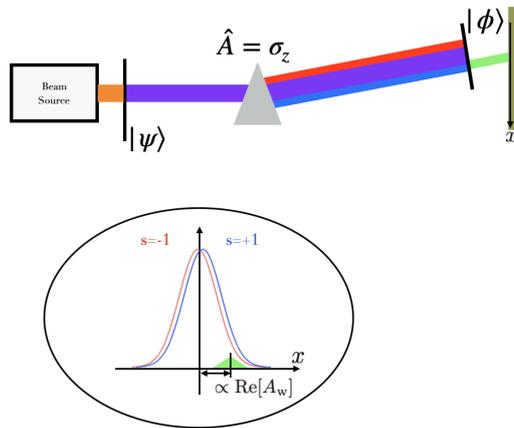


Figure 3.2: Depending on the choice of the preselected state $|\psi\rangle$ and the postselected state $|\phi\rangle$, the output of the measurement on the pointer variable appears selectively on outliers.

The following results are obtained by focusing on the real part and the imaginary part of θ in the first order of θ ,

$$2\text{Re}[\theta C_w] = 2\text{Re}[\theta]\text{Re}[C_w] - 2\text{Im}[\theta]\text{Im}[C_w]. \quad (3.45)$$

We can find from (3.44) that we can extract the weak value from the change in the postselection rate.

Another scheme using the lifetime of the particle has been proposed by Shomroni *et al.* [61] and us [55]. The parameter θ in (3.43) is time in this scheme and weak values are not obtained by a linear response of the postselection rate but by the mean lifetime. This scheme is also free from utilizing the meter system, also we can give another interpretation in which the scheme has a meter system. This scheme will be explained in detail in Chapter 5.

3.3.4 Preceding Experiments

Many experiments involving weak values have been conducted immediately after the theoretical proposal of weak values. The oldest result was published in 1991 [14].

One of the reasons why research on the amplification using the weak values attracted a great deal of attention was that Hosten and Kwiat succeeded in detecting the spin Hall effect of light (SHEL) using the WVA. They argued that *the measurement technique holds promise for precision metrology* in their paper [18]. Following the success of this SHEL experiment, Dixon *et al.* performed an ultrasensitive beam deflection (USBD) experiment in order to study the WVA itself to make its amplification effect more widely applicable. They cited Hosten and Kwiat's SHEL experiment as a successful example. The phrase 'weak value amplification' is coined in their paper.

In this subsection, we first explain the SHEL and USBD experiments. We then introduce some important experiments conducted in recent years and examples of experimental proposals.

Hosten's SHEL detection experiment

The spin Hall effect of light (SHEL) was proposed by Onoda, Murakami, and Nagaosa in 2004 [62]. The authors considered an equation of motion for photons in isotropic,

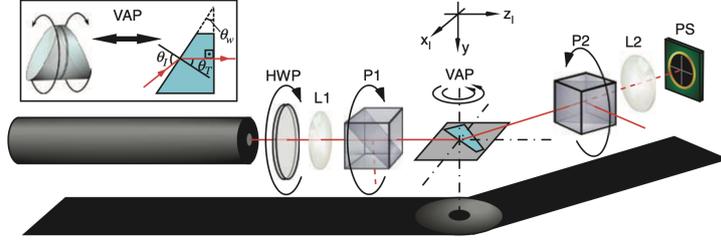


Figure 3.3: Fig.3 from [18]. Experimental setup of the SHEL experiment. In this experiment, the photon source is a laser that is depicted as a black bar in this figure. The spin Hall effect causes at the surface of the prism (VAP in this figure). Each preselected state $|\psi\rangle$ and postselected state $|\phi\rangle$ is selected by corresponding polarizers P1 and P2, respectively.

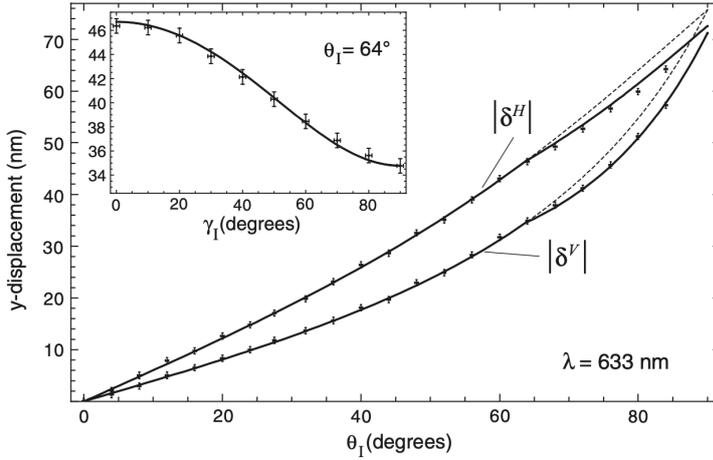


Figure 3.4: Fig.4 from [18]. The error bars are extremely small, and it seems that the measurement is realized with extremely high accuracy.

nonmagnetic medium which has the refractive index $n(\vec{r})$ that is not spatially uniform. They find a Berry-phase, which induces the SHEL effect causing a transverse shift of transmitted and reflected wave packets. However, no experiments at that time showed this shift in transmitted light.

They also predicted that such a shift would occur about 100 times the lattice constant of the medium. The lattice constant takes different values depending on the medium but is typically on the order of Angstrom. It means that to detect the SHEL effect, the shift is required on the order of 100 angstroms, that is, about 10 nanometers.

Although it is difficult to measure shifts on the order of 10 nanometers directly, they have used a preselection and postselection technique on the spin state to enhance the sensitivity. They insisted that they have attained sensitivity up to displacements of ~ 1 angstrom.

Dixon's USBD detection experiment

Dixon *et al.* have shown that tiny beam deflections and their corresponding angular deflections artificially made by a mirror can be detected using the WVA [19] to study the WVA itself. They prepared a Sagnac-type interferometer, as shown on the left side of Figure 3.5. One of the mirrors that make up this interferometer ('Piezo Driven Mirror' upper right, on the left of Figure 3.5) can be tilted. If this mirror is tilted, the beam's

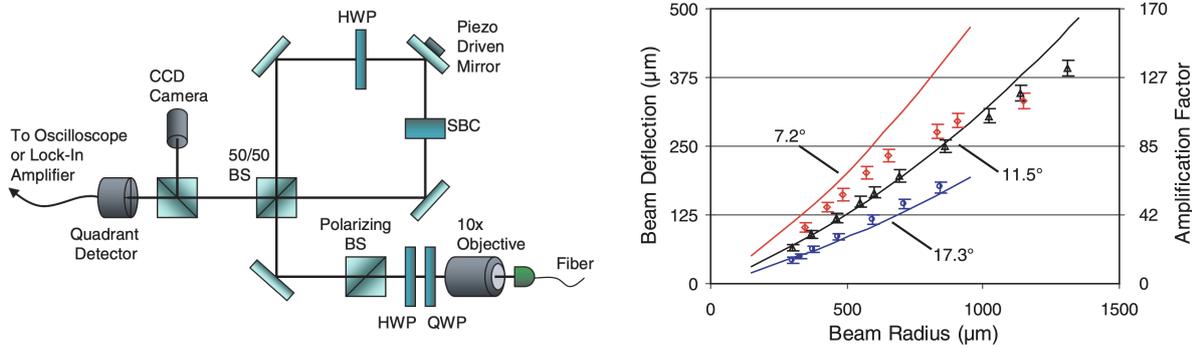


Figure 3.5: Left: Fig. 1 from [19], experimental setup of the USBD detection experiment. Right: Fig. 2 from [19], results of the experiment. The solid line is the shift of the theoretical prediction, and the dots are the experimental results. Each line corresponds to a change in the weak value. In this experiment, the beam radius can be adjusted by adjusting the lens, and the change in shift due to this adjustment is also measured.

direction changes, and the beam’s position shifts when it reaches the detector. Geometrical optics gives a prediction of this shift in the arrival position on the detector. They have fixed the mirror’s tilt, making the expected shift 2.95 micrometers.

In their experiment, the beamwidth is on the order of 100 micrometers. This larger beamwidth enables each component that passes clockwise and that passes counterclockwise through the interferometer to interfere with each other when the beam returns to the beam splitter. This interference changes the beam from the value expected by geometrical optics. We can understand this beam shift as weak value amplification through the quantum-classical correspondence seen under the paraxial approximation. In the paper [19], the beam’s actual shift detected at the detector occurs on the order of 100 micrometers (Figure 3.5).

They called this amplification in the sense that the original change of about three micrometers was able to be enlarged to the order of hundred micrometers.

Other experiments and proposals

After these studies, various researchers have performed many experiments using the WVA method. The goals of these experiments has also been diverse, ranging from small modifications of existing experiments to new applications. It is worth mentioning that some of these studies require different theoretical models to explain them, and discussions of such theoretical models have become intense.

Shomroni *et al.* demonstrated a weak measurement based on the dynamics of spontaneous emission of Rb atom in 2013 [61]. The spontaneous emission looks an exponential decay so that the excited Rb has a proper ‘lifetime’. However, they pointed out that the postselection on the polarization of the emitted photons enhances the mean lifetime of the excited Rb effectively (see Figure 3.7). The weak value is found in the extension of the mean lifetime of the spontaneous emission of Rb atoms.

On the other hand, Liping Xu *et al.* have measured the optical rotation change (2.73×10^{-4} rad) with a resolution of 6.75×10^{-7} rad and refractive index of chiral enantiomers such as sugars and amino acids [28]. In this experiment, the method proposed by Lan Luo in 2019 [42], which does not have the meter system, has been used. This study has the purpose of applying the WVA to the research of chemistry and hence can expand the application range of the WVA. In addition, there have been already many

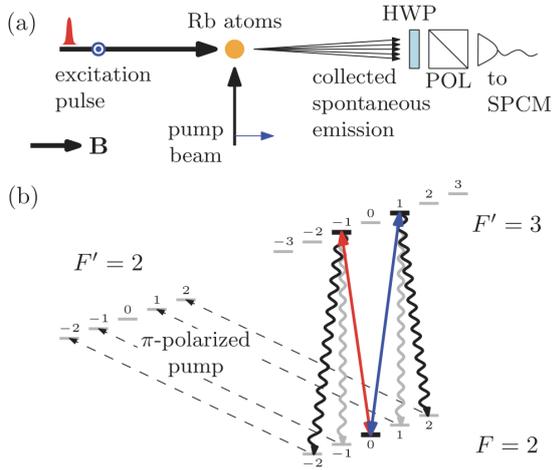


Figure 3.6: Fig 2. from [61] (a) Experimental setup. (b) Energy level of the Rb atoms in the experiment.

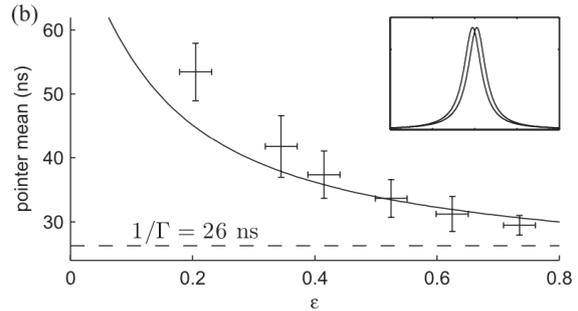


Figure 3.7: Fig 3. from [61]. Mean lifetime depends on the parameter ϵ , which determines the imaginary part of the weak value.

examples of various applied research.

Although most of the experiments involving the WVA are performed on the optical systems, other physical systems, such as neutrons [50, 51], are used in some experiments. In recent years, we have seen several proposals to apply the WVA in various systems. An example is found in quasiparticles in solid. Chen *et al.* have expected that the spin-dependent splitting induced by the spin-orbit interaction is related to the separation of Weyl nodes, a hallmark of a Weyl semimetal. They have also proposed that the distance of the nodes can be precisely determined using the WVA [63]. Besides, we also have proposals using cold atoms. Kawana and Ueda have proposed an application to the precise measurement of the gravitational constant as a method of using cold atoms [46].

There are still new proposals in the study of optical systems. For example, Carrasco and Orszag have proposed weak measurements to enlarge the radiation pressure effect of a single-photon on a mechanical oscillator placed in the middle of a Fabry-Pérot cavity and initialized in the ground state [64].

3.4 Controversial points of the WVA

So far, we have discussed the basic theory of the WVA and some successful examples of experiments. We now mention several unwanted points on the WVA. Such unwanted points have been pointed out early on. In fact, Dixon *et al.* have already mentioned that

like any amplifier, something must be sacrificed in order to achieve the enhancement of the signal. For weak values the sacrifice comes in the form of throwing away most of the data in the post-selection process [19].

We can classify these unwanted points into two categories. The first is concerned with the problem of decreasing the data in the postselection. The second, which has not been mentioned in the quote above, is that postselection causes uncertainty related to the state of the meter system e intricately in the measurement results. Although such uncertainty has already appeared in the above-mentioned first-order approximation, it can become much more complicated when the higher-order terms are examined.

3.4.1 Issues involving the decreasing of the number

In the WVA, we can use only the amount of data may be reduced considerably due to the postselection. This results in two undesirable points. One is that the rate of noise increases as the signal decreases. The other is that statistical accuracy is expected to deteriorate.

Vulnerability to contamination with incorrect states

So far, we assume that we can always prepare the pure initial state $|\psi\rangle$. However, there is no guarantee that one can choose $|\psi\rangle$ perfectly. To discuss such uncertainty in the state preparation, let us assume that the actual initial state is not $|\psi\rangle$ (or its density matrix representation $|\psi\rangle\langle\psi|$) but

$$\rho_{\text{ini}} := (1 - \epsilon)|\psi\rangle\langle\psi| + \epsilon|\psi^\perp\rangle\langle\psi^\perp|, \quad (3.46)$$

where $|\psi^\perp\rangle$ is orthogonal to the state $|\psi\rangle$.

For this mixed state, the probability on the final measurement of the pointer variable x without the postselection $P_{\rho_{\text{ini}}}(x)$ is

$$\begin{aligned} P_{\rho_{\text{ini}}}(x) &= \text{tr} \left[\hat{1} \otimes |x\rangle\langle x| \hat{U}(\theta) \rho_{\text{ini}} \otimes |\chi\rangle\langle\chi| \hat{U}^\dagger(\theta) \right] \\ &= (1 - \epsilon) \text{tr} \left[\hat{1} \otimes |x\rangle\langle x| \hat{U}(\theta) |\psi\rangle\langle\psi| \otimes |\chi\rangle\langle\chi| \hat{U}^\dagger(\theta) \right] \\ &\quad + \epsilon \text{tr} \left[\hat{1} \otimes |x\rangle\langle x| \hat{U}(\theta) |\psi^\perp\rangle\langle\psi^\perp| \otimes |\chi\rangle\langle\chi| \hat{U}^\dagger(\theta) \right] \\ &= (1 - \epsilon) P_{|\psi\rangle}(x) + \epsilon P_{|\psi^\perp\rangle}(x). \end{aligned} \quad (3.47)$$

Here,

$$P_{|\psi\rangle}(x) := |\langle x | \hat{U}(\theta) | \psi \rangle | \chi \rangle|^2, \quad (3.48)$$

is the probability distribution of the pointer variable when the initial state of the target system is selected as $|\psi\rangle$, and

$$P_{|\psi^\perp\rangle}(x) := |\langle x | \hat{U}(\theta) | \psi^\perp \rangle | \chi \rangle|^2, \quad (3.49)$$

is the probability distribution of the pointer variable when the initial state of the target system is selected as $|\psi^\perp\rangle$. The mixture ratio can be considered as a SN ratio

$$\text{SN}_{\text{st}} := \frac{(1 - \epsilon)}{\epsilon}. \quad (3.50)$$

On the other hand, when the postselection is performed, the SN ratio should be different from (3.50). The probability on the final measurement of the pointer variable x after succeeding the postselection (the postselected state is $|\phi\rangle$) $P(x, \rho_{\text{ini}} \rightarrow |\phi\rangle\langle\phi|)$ is

$$\begin{aligned} P(x, \rho_{\text{ini}} \rightarrow |\phi\rangle\langle\phi|) &= \langle\phi| \langle x | \hat{U}(\theta) \rho_{\text{ini}} \otimes |\chi\rangle\langle\chi| \hat{U}^\dagger(\theta) | \phi \rangle | x \rangle \\ &= (1 - \epsilon) |\langle\phi| \langle x | \hat{U}(\theta) | \psi \rangle | \chi \rangle|^2 + \epsilon |\langle\phi| \langle x | \hat{U}(\theta) | \psi^\perp \rangle | \chi \rangle|^2. \end{aligned} \quad (3.51)$$

Meanwhile, the square of the absolute value of the amplitude (3.23),

$$P(x | \psi \rightarrow |\phi\rangle) := |\langle x | \xi \rangle|^2 = \frac{|\langle\phi| \langle x | \hat{U}(\theta) | \psi \rangle | \chi \rangle|^2}{|\langle\phi| \hat{U}(\theta) | \psi \rangle | \chi \rangle|^2}, \quad (3.52)$$

gives a conditional probability distribution that the postselection is successfully performed. Using (3.52), the probability distribution (3.51) can be rewritten as

$$\begin{aligned} P(x, \rho_{\text{ini}} \rightarrow |\phi\rangle\langle\phi|) \\ = (1 - \epsilon) |\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|^2 P(x|\psi \rightarrow |\phi\rangle) + \epsilon |\langle\phi|\hat{U}(\theta)|\psi^\perp\rangle|\chi\rangle|^2 P(x|\psi^\perp \rightarrow |\phi\rangle). \end{aligned} \quad (3.53)$$

One may obtain a conditional probability distribution $P(x|\rho_{\text{ini}} \rightarrow |\phi\rangle\langle\phi|)$ from (3.53) by introducing a proper normalization constant N as

$$\begin{aligned} P(x|\rho_{\text{ini}} \rightarrow |\phi\rangle\langle\phi|) \\ = (1 - \epsilon) \frac{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|^2}{N} P(x|\psi \rightarrow |\phi\rangle) + \epsilon \frac{|\langle\phi|\hat{U}(\theta)|\psi^\perp\rangle|\chi\rangle|^2}{N} P(x|\psi^\perp \rightarrow |\phi\rangle). \end{aligned} \quad (3.54)$$

The first term is the signal obtained through an appropriate process, while the second term is derived from noise. The conditional probability distribution (3.54) that meets the conditions given in preselection and postselection can be described as a stochastic mixture by adopting the mixture ratio of the obtained probability distributions as the SN, as in the case of no postselection,

$$\text{SN}_{\text{p.s.}} = \frac{(1 - \epsilon) |\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|^2}{\epsilon |\langle\phi|\hat{U}(\theta)|\psi^\perp\rangle|\chi\rangle|^2}, \quad (3.55)$$

$$= \frac{|\langle\phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle|^2}{|\langle\phi|\hat{U}(\theta)|\psi^\perp\rangle|\chi\rangle|^2} \text{SN}_{\text{st.}} \quad (3.56)$$

In the weak limit, (3.56) becomes

$$\text{SN}_{\text{p.s.}} = \frac{|\langle\phi|\psi\rangle|^2}{|\langle\phi|\psi^\perp\rangle|^2} \text{SN}_{\text{st.}} \quad (3.57)$$

To use the weak value as an amplifier, we must choose the preselected state $|\psi\rangle$ and the postselected state $|\phi\rangle$ to realize a large weak value A_w . However, as can be seen from (3.36) and (3.37), for the weak value A_w to take a large value, it is necessary to select the postselected state $|\phi\rangle$ almost orthogonal to the preselected state $|\psi\rangle$. This implies that $|\langle\phi|\psi\rangle|^2$ must become smaller and the SN must deteriorate.

This result puts a practical limit on the use of the WVA. In particular, when the target system is a two-level system, the SN ratio is found to be

$$\begin{aligned} \text{SN}_{\text{p.s.}} &= \frac{|\langle\phi|\psi\rangle|^2}{1 - |\langle\phi|\psi\rangle|^2} \text{SN}_{\text{st.}}, \\ &= \frac{|\langle\phi|\psi\rangle|^2}{1 - |\langle\phi|\psi\rangle|^2} \frac{1 - \epsilon}{\epsilon}. \end{aligned} \quad (3.58)$$

Because both ϵ and $|\langle\phi|\psi\rangle|^2$ are considered to be much smaller than 1 in a realistic situation, the SN ratio becomes

$$\frac{|\langle\phi|\psi\rangle|^2}{\epsilon}, \quad (3.59)$$

using $1/(1 - |\langle\phi|\psi\rangle|^2) \approx 1 + |\langle\phi|\psi\rangle|^2$. To obtain a meaningful result, the postselection rate should be larger than the rate of noise contamination ϵ . Unless the noise can also be

reduced as well as the signal, we should be careful of this SN ratio not being deteriorated too much.

It seems that this problem has not been picked up in many papers, but it seems to be important in actual experiments. An example of a quantitative analysis of this problem is a study by Ueda and Kitahara which applied weak measurement to neutron EDM measurement [47]. Dixon *et al.* also mentioned a similar problem in their paper [19], although no quantitative discussion was made.

Statistical weakness

Even if the above noise contamination is completely prevented, the weaknesses of discarding the signals remain. Since the quantum theory is based on the probability theory, this problem is unavoidable from the view point of statistical accuracy. If the change given to the distribution by the postselection becomes larger, the amount of the information on the parameter we want to estimate obtained experimentally becomes larger. However, we can improve statistical accuracy by collecting a large number of statistics.

When we consider the simple WVA model and overlook various problems in the actual measurements, the accuracy of the WVA experiment cannot surpass the accuracy of the corresponding experiment without the postselection in view of the parameter estimation theory [30].

Other comments on this issue

The above points are undesirable as these are usually essential on the precision measurements. Therefore, researches related to these points have been conducted in various respects. In recent years, some researchers have studied photon recycling to solve this issue. The idea is that the experimenter collects and reuses photons that failed to pass the postselection [38, 65].

It is also known that in actual measurement, it is not always the case to increase the number of statistics. As an example, the paper [66] discusses saturation caused by the high sensitivity of the detector.

3.4.2 Dependence of the meter state

We find in (3.32) that if we only use the first order of θ , the shift of the pointer variable $\Delta_w(\theta)$ depends on the state of the meter system when $\text{Im}[A_w]$ is nonzero. Thus, if we cannot restrict the preselected and postselected states satisfying the condition $\text{Im}[A_w]$ is zero, some uncertainty will occur. Such uncertainty is also caused by higher-order terms that are ignored in the first-order approximation. This uncertainty can be reduced by setting the model more accurately, making computational assumptions more precisely, eliminating approximations.

In addition, the effects of these higher-order terms typically work to counteract the amplification effect of increasing the weak value. As we show in Chapter 4, this point provides a limit of the effectiveness of the WVA.

3.5 Technical points in actual experiments

So far, we have discussed the WVA and its related issues using a simple theoretical model of measurement. When we apply the idea of the WVA, we need to consider some technical points that have not been mentioned in the above. These technical points

have been addressed individually for the system used in each experiment. Nonetheless, there are common issues in the optical experiments analyzed below and the high energy experiment discussed in Chapter 5.

3.5.1 Setting of target system and meter system in experiments using photons

Photons are particles that emerge from the degrees of freedom of the vector field $A_\mu(x)$ derived from the Lagrangian that has $U(1)$ gauge symmetry. By properly considering the solution of the equation describing the vector field after the quantization, the field operator is given by

$$\hat{A}_\mu(x) = \int \frac{d^3p}{(2\pi)^3 \sqrt{2E_{\mathbf{p}}}} \sum_{r=0}^3 (\hat{a}_{\mathbf{p}}^r \epsilon_\mu^r(p) e^{-ip \cdot x} + \hat{a}_{\mathbf{p}}^{r\dagger} \epsilon_\mu^{r*}(p) e^{ip \cdot x}), \quad (3.60)$$

where $\hat{a}_{\mathbf{p}}^r$ and $\hat{a}_{\mathbf{p}}^{r\dagger}$ are creation and annihilation operators satisfying

$$[\hat{a}_{\mathbf{p}}^r, \hat{a}_{\mathbf{q}}^{s\dagger}] = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q}), \quad (3.61)$$

and ϵ^r is polarization vectors [2]. Then, an arbitrary single-photon state $|\Psi\rangle$ corresponding to such a field is described as

$$|\Psi\rangle := \sum_{s=0}^3 \int \frac{d^3p}{\sqrt{(2\pi)^3 2E_{\mathbf{p}}}} \sqrt{2E_{\mathbf{p}}} \Psi(s, \mathbf{p}) \hat{a}_{\mathbf{p}}^{s\dagger} |0\rangle, \quad (3.62)$$

$$= \sum_{s'=1,2} \int \frac{d^3p}{\sqrt{(2\pi)^3}} \Psi(s', \mathbf{p}) \hat{a}_{\mathbf{p}}^{s'\dagger} |0\rangle, \quad (3.63)$$

where $|0\rangle$ is the vacuum state satisfying $\hat{a}_{\mathbf{p}}^s |0\rangle = 0$ for all (s, \mathbf{p}) . Since the polarization is a vector defined in space-time, the label s is taken from zero to three, but it can be limited to a two-dimensional subspace due to the gauge symmetry and the equation of motion. Accordingly, the label is now restricted to $s' = 1, 2$ in (3.63). $\Psi(s, \mathbf{p})$ and $\Psi(s', \mathbf{p})$ are wave functions.

The polarization and the momentum are the degrees of freedom possessed by the photon, but unlike position and momentum, or spins in different directions, they can be measured simultaneously. Time evolutions of the spin and momentum are independent, as long as no physical interaction between them is given. That means that they can be regarded as independent physical systems, and we can consider the polarization and the momentum as different physical systems and can use them as a target system \mathcal{H} and a meter system \mathcal{K} , respectively.

In our discussion, we prepare the following product of the functions as the initial state in the actual measurement.

$$\Psi(s', \mathbf{p}) = c_{s'} \chi(\mathbf{p}) \quad (3.64)$$

Using the state (3.64), we can check that there is no correlation between the polarization and the momentum so that we can write the state of the polarization as

$$|\psi\rangle := c_1 |1\rangle + c_2 |2\rangle, \quad (3.65)$$

where $|1\rangle$ and $|2\rangle$ are one particle states representing the polarization $s' = 1$ and $s' = 2$ respectively. Also, the momentum is represented by

$$|\chi\rangle := \int d^3p \chi(\mathbf{p}) |\mathbf{p}\rangle, \quad (3.66)$$

where $|\mathbf{p}\rangle$ is the momentum eigenstate.

Finally, we mention the precautions when applying the WVA method to the particle experiment, which is related to the present discussion. In the theory dealing with elementary particles, for each particle, a quantum-theoretical description is obtained by using the corresponding field theory with quantization process. From this viewpoint, we should refer to the field theory on the degree of freedom that the particle has. However, it is not always necessary to treat all degrees of freedom in our discussions. For example, although the photon has the degrees of freedom of polarization and those of energy and momentum, we do not always need to consider them at the same time. In fact, we often encounter the creation and annihilation operators in textbooks (*e.g.* [67]) and papers (*e.g.* [68]) which have no label for polarization and momentum or either one of them. In such cases, these creation and annihilation operators are basically considered to be composed of integral operations. Explicitly, the ‘reduced’ annihilation operator defined by

$$\hat{a} := \sum_{s'=1,2} \int \frac{d^3p}{\sqrt{(2\pi)^3}} \Psi(s', \mathbf{p}) \hat{a}_{\mathbf{p}}^{s'}, \quad (3.67)$$

where $\Psi(s', \mathbf{p})$ is a wave function used in (3.63), satisfies the commutation relation,

$$\begin{aligned} [\hat{a}, \hat{a}^\dagger] &= \sum_{r,s} \int \frac{d^3p}{\sqrt{(2\pi)^3}} \frac{d^3q}{\sqrt{(2\pi)^3}} \Psi^*(r, \mathbf{p}) \Psi(s, \mathbf{q}) [\hat{a}_{\mathbf{p}}^r, \hat{a}_{\mathbf{q}}^{s\dagger}] \\ &= \sum_s \int d^3p |\Psi(s, \mathbf{p})|^2 = 1. \end{aligned} \quad (3.68)$$

When the separable condition (3.64) is satisfied, we can also define ‘partially reduced’ annihilation operators,

$$\hat{a}^r := \int \frac{d^3p}{\sqrt{(2\pi)^3}} \chi^*(\mathbf{p}) \hat{a}_{\mathbf{p}}^r, \quad (3.69)$$

$$\hat{a}_{\mathbf{p}} := \sum_{s'=1,2} c_{s'}^* \hat{a}_{\mathbf{p}}^{s'}, \quad (3.70)$$

satisfying the expected commutation relation of the creation and annihilation operators,

$$\begin{aligned} [\hat{a}^r, \hat{a}^{s\dagger}] &= \int \frac{d^3p}{\sqrt{(2\pi)^3}} \frac{d^3q}{\sqrt{(2\pi)^3}} \chi^*(\mathbf{p}) \chi(\mathbf{q}) [\hat{a}_{\mathbf{p}}^r, \hat{a}_{\mathbf{q}}^{s\dagger}], \\ &= \int d^3p \delta^{rs} |\chi(\mathbf{p})|^2, \\ &= \delta^{rs}, \end{aligned} \quad (3.71)$$

and

$$\begin{aligned} [\hat{a}_{\mathbf{p}}, \hat{a}_{\mathbf{q}}^\dagger] &= \sum_{r,s=1,2} c_r^* c_s [\hat{a}_{\mathbf{p}}^r, \hat{a}_{\mathbf{q}}^{s\dagger}] \\ &= (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \sum_{r=1,2} |c_r|^2 \\ &= (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}). \end{aligned} \quad (3.72)$$

When (3.63) is applicable, we can ignore the momentum degrees of freedom in the discussions, which we will do in Chapter 5 when we perform calculations limited to polarization and flavor degrees of freedom. On the other hand, if the momentum of the photon and its spin are entangled such that we cannot use (3.64), the reduced operator (3.69) cannot be defined uniquely. In order to use the flavor degrees of freedom of K^* mesons as a two-level system in the same way as the polarization degrees of freedom of photons discussed above, we must assume a similar condition (3.64) for the other degrees of freedom of K^* mesons.

3.5.2 Points to consider common to optical experiments

In the above discussion, we have not considered the time evolution from the initial state to before the interaction to obtain the state (3.5). However, the states of the target system and the meter system prepared in the actual experiment may evolve, since they may not necessarily be the eigenstates of the Hamiltonian. Further, especially in an experiment using a photonic system, experimenters can change the state using optical elements. We should take these elements into the theoretical model as well.

For these cases, both the state of the target system $|\psi\rangle$ and $|\chi\rangle$ should be regarded as states just before the interaction $\hat{U}(\theta)$ for (3.5). Also, $\langle\phi|$ and $\langle x|$ should be regarded states just after the interaction $\hat{U}(\theta)$ for (3.18) and (3.23). There may be a certain amount of time evolution between the operation performed by the experimenter and the interaction, and the insertion of optical elements may occur additionally. Therefore, it is necessary to take these elements into consideration when making corrections.

In our cases, fortunately, the state of the target system can only be changed by the preselection and the postselection and the time evolution of the target system does not occur. Due to this reason, the state of the target system is affected at only three stages, the preselection, the interaction, and the postselection. Due to the commutativity of these time evolution operators, it can be considered that the states immediately before and after the interaction can be selected for the target system without considering time evolution.

On the other hand, the state of the meter system is affected by free time evolution and insertion of the lens. To distinguish between these changes in state, we define the following states: Let the initially prepared meter state be $|\chi_{pre}\rangle$. Also, let the state we finally measure the pointer variable be $|\xi_m\rangle$.

Using the results of Fourier optics, we can specify the operator describing the free time evolution of the photon and the operator representing the effect of passing the lens. We discuss the details of its derivation in Appendix A. Referring to these results, the operator of free evolution $\hat{V}(z)$ and the operator representing the action of a lens \hat{V}_L are given by

$$\hat{V}(z) = e^{-\frac{i}{2k_0}z\hat{P}^2}, \quad (3.73)$$

$$\hat{V}_L = e^{-i\frac{k_0}{2l_f}\hat{X}^2}, \quad (3.74)$$

where z is the distance traveled along the optical axis, k_0 is the momentum of the photon (*i.e.* wave number) and l_f is the focal length of the lens.

In either the SHEL experiment or the USBD experiment, the state just after the interaction $|\xi\rangle$ is transformed into $|\xi_m\rangle$ by the free time evolution (see Figures 3.8 and 3.9). Although lens 2 is inserted before measuring the pointer variable in the SHEL experiment, the purpose of this insertion is to substantially eliminate the time evolution of the photon after passing through the lens and also to fix the state of the meter system.

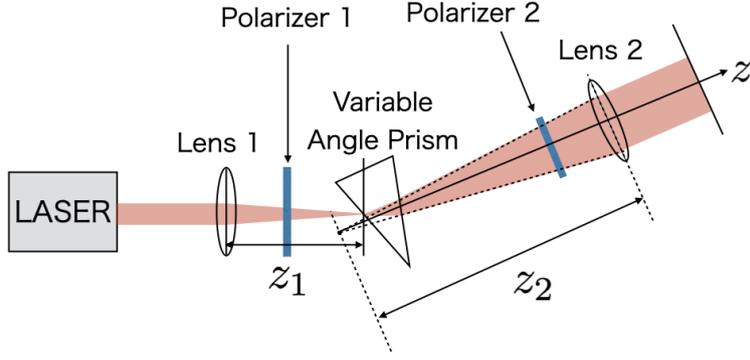


Figure 3.8: Schematic diagram of the SHEL experiment. The interaction between the target system (spin of the photon) and the meter system (position at the screen right end this figure) occurs at the surface of the variable angle prism.

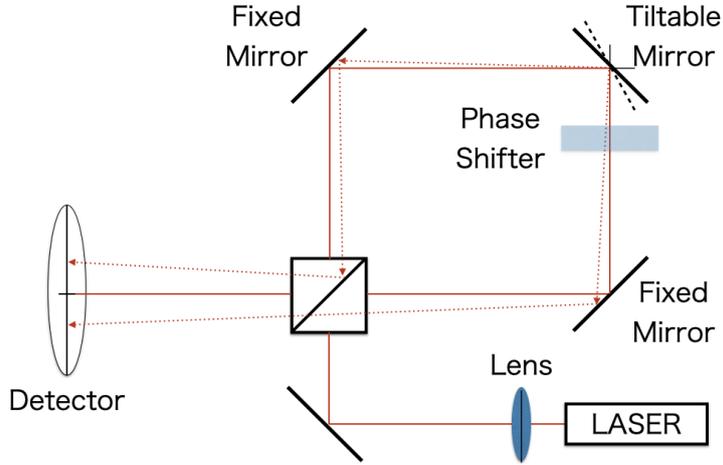


Figure 3.9: The setup of the USBD experiment, with a slight modification for the convenience of our argument.

This means that the state $|\xi_m\rangle$ is given by

$$|\xi_m\rangle = \hat{V}(z) |\xi\rangle, \quad (3.75)$$

where z is the distance between where the interaction occurs and the location where we measure the pointer variable. Using (3.75), we obtain the expectation value of the pointer variable,

$$\langle \xi_m | \hat{X} | \xi_m \rangle = \langle \xi | \hat{V}^\dagger(z) \hat{X} \hat{V}(z) | \xi \rangle, \quad (3.76)$$

and the variance of the pointer variable,

$$\langle \xi_m | \hat{X}^2 | \xi_m \rangle - (\langle \xi_m | \hat{X} | \xi_m \rangle)^2 = \langle \xi | \hat{V}^\dagger(z) \hat{X}^2 \hat{V}(z) | \xi \rangle - (\langle \xi | \hat{V}^\dagger(z) \hat{X} \hat{V}(z) | \xi \rangle)^2. \quad (3.77)$$

To calculate (3.76) and (3.77), the following relations are useful,

$$\hat{V}^\dagger(z) \hat{X} \hat{V}(z) = \hat{V}^\dagger(z) \left[\hat{X}, \hat{V}(z) \right] + \hat{X} = \hat{X} + \frac{z}{k_0} \hat{P} \quad (3.78)$$

$$\begin{aligned} \hat{V}^\dagger(z) \hat{X}^2 \hat{V}(z) &= \hat{V}^\dagger(z) \hat{X} \hat{V}(z) \hat{V}^\dagger(z) \hat{X} \hat{V}(z) \\ &= \hat{X}^2 + \frac{z}{k_0} \left\{ \hat{X}, \hat{P} \right\} + \frac{z^2}{k_0^2} \hat{P}^2. \end{aligned} \quad (3.79)$$

In fact, by making good use of (3.78) and (3.79) in future calculations, it is not necessary to explicitly calculate the state $|\xi_m\rangle$ to reach the results.

For our convenience, we write a useful relation to be used in our analysis. If the observable of the target system satisfies the condition $\hat{A}^2 = \hat{1}$, we have

$$\begin{aligned}
\langle \phi | U(\theta) | \psi \rangle | \chi \rangle &= \langle \phi | e^{-i\theta \hat{A} \otimes \hat{Y}} | \psi \rangle | \chi \rangle \\
&= \sum_{k=0}^{\infty} \frac{(-i\theta)^k}{k!} \langle \phi | \hat{A}^k \otimes \hat{Y}^k | \psi \rangle | \chi \rangle \\
&= \sum_{k=0}^{\infty} \frac{(-\theta^2)^k}{2k!} \langle \phi | \hat{1} \otimes \hat{Y}^{2k} | \psi \rangle | \chi \rangle - i \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} \langle \phi | \hat{A} \otimes \hat{Y}^{2k} | \psi \rangle | \chi \rangle \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k}}{2k!} \langle \phi | \psi \rangle \hat{Y}^{2k} | \chi \rangle - i \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} \langle \phi | \hat{A} | \psi \rangle \hat{Y}^{2k} | \chi \rangle \\
&= \langle \phi | \psi \rangle \left[\cos(\theta \hat{Y}) - i A_w \sin(\theta \hat{Y}) \right] | \chi \rangle.
\end{aligned} \tag{3.80}$$

This result (3.80) can be used for any two level systems as long as $\hat{A}^2 = \hat{1}$ holds.

3.5.3 Examples of state selection in the target system

We discuss how to perform the preselection and postselection using two examples, the SHEL experiment and the USBD experiment here.

SHEL experiment

In the SHEL experiment, the spin of the photon is the target system. It is known that the eigenstates of photon spins correspond to the circularly polarized states (see, *e.g.* [69]). In optical systems, we can select the state of the polarization using polarizers, wave-plates, etc. The preselected state is given by polarizer 1, and the postselected state is given by polarizer 2 (see Figures 3.8 and 3.10).

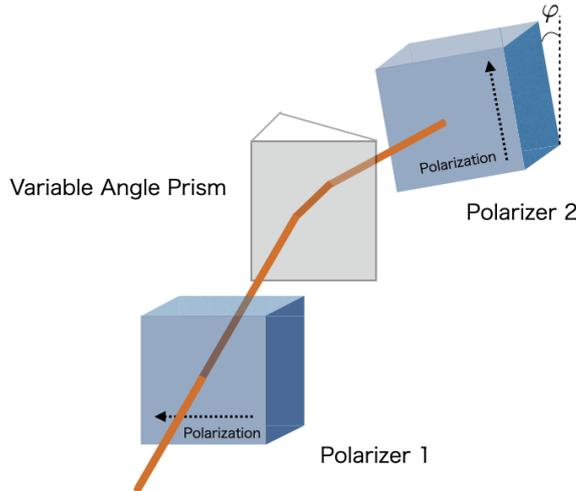


Figure 3.10: Preselection by Polarizer 1 and postselection by Polarizer 2 tilted by angle φ .

The polarization state that can be specified using these polarizers is a linearly polarized light, described as a superposition of clockwise and counterclockwise circular polarization.

The preselected state $|\psi\rangle$ is given by

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle), \quad (3.81)$$

where $|+\rangle$ is the spin 1 state and $|-\rangle$ is the spin -1 state. Also, the postselected state is given by

$$|\phi\rangle = \frac{1}{\sqrt{2}i} (e^{i\varphi} |+\rangle - e^{-i\varphi} |-\rangle), \quad (3.82)$$

where φ can be given by adjusting the angle parameter φ of the polarizer 2 (see Figure 3.10).

The weak value, which is defined in (3.21), then reads

$$A_w = i \cot \varphi. \quad (3.83)$$

This is purely imaginary.

USBD experiment

In the USBD experiment, the experimenters employed an indirect measurement to find ‘whether the photon undergone clockwise or counterclockwise’ when it went through the Sagnac interferometer. To achieve this, they set a tiltable mirror in the experimental setup. (see Figure 3.5 or its simplified version Figure 3.9) The tilt of the mirror caused a different shift of the beam direction depending on the route clockwise or counterclockwise. In this case, the target system is a degree of freedom describing whether the photon undergoes clockwise or counterclockwise path in the interferometer.

We express the state corresponding to the photon going in the clockwise direction as $|r\rangle$, and the state corresponding to the photon going in the counterclockwise direction as $|l\rangle$. We assume that they form an orthonormal basis of the target system.

We construct the target observable \hat{A} corresponding to ‘whether the photon undergone clockwise or counterclockwise path’ as follows

$$\hat{A} = |r\rangle \langle r| - |l\rangle \langle l|. \quad (3.84)$$

To perform the preselection and postselection, the beam splitter and the phase shifter are used in Figure 3.9. We explain them here.

Beam splitter In the beam splitter, a relative phase is given between the transmission component and the reflection component. In this experiment, the beam splitter can be represented by the operator M_{BS} defined by (Figure 3.11),

$$\hat{M}_{BS} |b\rangle = \frac{1}{\sqrt{2}}(|r\rangle + i |l\rangle), \quad (3.85)$$

$$\hat{M}_{BS} |d\rangle = \frac{1}{\sqrt{2}}(i |r\rangle + |l\rangle). \quad (3.86)$$

Since the beam from the light source is incident from the ‘bright port’ of the beam splitter, it designates a counterclockwise and clockwise superposition state (3.85).

In this experiment, due to the structure of the interferometer, the beam output $|r\rangle$ returns to the port corresponding to the output $|l\rangle$, and the beam output $|l\rangle$ returns to the port corresponding to the output $|r\rangle$. When the phase shifter is removed, the beam

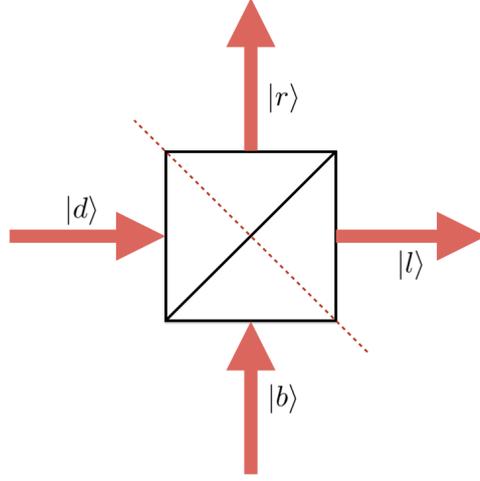


Figure 3.11: The beam splitter gives a relation between an input/output relation.

returns to the originally incident port (bright port), so the output from the bright port can be understood as

$$\frac{1}{\sqrt{2}}(|r\rangle + i|l\rangle). \quad (3.87)$$

Similarly, the state corresponding to the output from the dark port is

$$\frac{1}{\sqrt{2}}(i|r\rangle + |l\rangle). \quad (3.88)$$

Therefore, we consider that the state of the target system is (3.88) when the photon goes into to the dark port.

Phase shifter In the USBD experiment, they use the half-wave plate (HWP) and the Soleil-Babinet compensator (SBC) (see Figure 3.5). When the vertically polarized state passes through the HWP, it changes to a polarized state orthogonal to the state before it passes through. The polarization state of the photon that goes clockwise changed through to a state orthogonal to the initial polarization state. At the time of reflecting the tiltable mirror, the photon that goes through the interferometer clockwise and the photon that goes through counterclockwise are orthogonal to each other in terms of the state of the polarization. In the mean time, the SBC gives a phase difference between the polarization of the photon. This phase difference is adjustable. By combining the HWP and the SBC, a phase shifter that gives a path-dependent phase difference is constructed. Figure 3.9 is a simplified description where only SBC is used as a phase shifter.

Due to the alignment of the mirror, SBC, and HWP, the order of the phase shift due to the passage of the phase shifter and the reflection at the mirror changes depending on the state of the photon. If we want to describe the state considering the difference in the order, the state of the photon at the stage of returning to the beam splitter should be described as

$$\hat{O}_l e^{-i\theta \hat{A} \otimes \hat{X}} \hat{O}_r M_{\text{bs}} |b\rangle, \quad (3.89)$$

where $e^{-i\theta \hat{A} \otimes \hat{X}}$ is an operator corresponding to the tiltable mirror. The operators \hat{O}_l and \hat{O}_r are given by

$$\hat{O}_r = e^{i\frac{\pi}{2}} |r\rangle \langle r| + |l\rangle \langle l|, \quad \hat{O}_l = |r\rangle \langle r| + e^{-i\frac{\pi}{2}} |l\rangle \langle l|, \quad (3.90)$$

which describe the phase shifter.

By the way, since the base that diagonalizes the three operators \hat{A} , \hat{O}_l , and \hat{O}_r is commonly given by the same set $\{|r\rangle$ and $|l\rangle\}$, they are all commutative with each other. This means that the state (3.89) is

$$e^{-i\theta\hat{A}\otimes\hat{X}}\hat{O}_l\hat{O}_rM_{\text{bs}}|b\rangle, \quad (3.91)$$

and the state $\hat{O}_l\hat{O}_rM_{\text{bs}}|b\rangle$ can be understood as the preselected state $|\psi\rangle$.

The weak value, which is defined in (3.21), then reads

$$\begin{aligned} A_w &= \frac{\langle d_{\text{out}}|\hat{A}|\hat{O}_l\hat{O}_rM_{\text{bs}}|b\rangle}{\langle d_{\text{out}}|\hat{O}_l\hat{O}_rM_{\text{bs}}|b\rangle} \\ &= \frac{\frac{1}{\sqrt{2}}(-i\langle r| + \langle l|) \times \hat{A} \times \frac{1}{\sqrt{2}}(e^{i\frac{\varphi}{2}}|r\rangle + ie^{-i\frac{\varphi}{2}}|l\rangle)}{\frac{1}{\sqrt{2}}(-i\langle r| + \langle l|) \times \frac{1}{\sqrt{2}}(e^{i\frac{\varphi}{2}}|r\rangle + ie^{-i\frac{\varphi}{2}}|l\rangle)} \\ &= \frac{-i \cos \frac{\varphi}{2}}{\sin \frac{\varphi}{2}} = -i \cot \frac{\varphi}{2}, \end{aligned} \quad (3.92)$$

which is purely imaginary.

Having provided the common bases of the two optical experiments, we next examine the validity of the two experiments concerning the WVA.

Chapter 4

Effects on the WVA

Lee and Tsutsui have proposed a scheme to evaluate the merit of the experiment involving the WVA in 2014 [36]. Since this scheme is based on the idea of measurement uncertainty, it has been expected to be applied to analyzing the results of actual experiments. Also, it has been expected to be used for judging the validity of the results of the experiments. To assure these points, we have analyzed the experimental results by applying the scheme to two reliable preceding experiments, one of which is Hosten's detection of the spin Hall effect of light (SHEL) [18] and the other is Dixon's ultrasensitive beam deflection (USBDF) experiment [19]. The results of these analyses have already been published [54].

In this chapter, Sections 4.1 and 4.2 are used for explaining the method of analysis. The method is based on the paper [36]. However, it lacks some of the perspectives needed to apply the method to the actual experiments and to perform an analysis (*e.g.*, time evolution shown in Subsection 3.5.2), and accordingly, the method is not strictly consistent with the model that describes the actual experiments. Besides, there were some technically unavoidable problems in applying their method to experiments and conducting analyses. In the paper [54], the present author has solved these technical issues and obtained the results shown in this chapter. (Those points and results are indicated when they appear below.) In Section 4.3, we introduce the experiments and corresponding calculation models for describing the analyzed experiments. Section 4.4 is devoted to explaining how to estimate the relative uncertainty in the actual calculation. Finally, we see the results in Section 4.5.

The present author's contribution in this study is to match the proposal [36] for analytical methods with the theory that describes the actual experiments and to apply the analytical method thus matched to the actual experiments. As a result, that proposal has been improved in view of the applicability to analyzing actual experiments.

4.1 Measurement Uncertainty

The measurement uncertainty gives an indicator of the validity of measurements in experiments and theoretical expectations. We consider that the measurement uncertainty should be significant in evaluating the performance of the WVA. Since we have considered the shift of the expectation value of the pointer variable in Chapter 3, we should be careful of the uncertainty of the expectation value.

In the measurement, suppose that we acquire n outputs from the detector by measuring the observable \hat{X} under the meter state $|\xi_m\rangle$ obtained after the postselection, and let $\{\tilde{x}_i,$

$i = 1, \dots, n\}$, be the values of the outputs. We then have the experimental average

$$\tilde{X}_n = \frac{1}{n} \sum_{i=1}^n \tilde{x}_i, \quad (4.1)$$

which is to be compared to the expectation value $\langle \xi_m | \hat{X} | \xi_m \rangle_{\text{app}}$ evaluated by using some assumptions and approximation methods in a given theoretical model of the total system.

Now, let us consider the difference between the theoretically obtained approximate expectation value and the experimentally obtained average value,

$$\epsilon_{\text{op}}(n) := \left| \langle \xi_m | \hat{X} | \xi_m \rangle_{\text{app}} - \tilde{X}_n \right|. \quad (4.2)$$

We note that this is an operationally meaningful quantity in the sense that it can be obtained directly from our experiment and the theory we use. In order to estimate the error (4.2), we classify the uncertainties.

First of all, we focus on each of the two components contained in (4.2), \tilde{X}_n and $\langle \xi_m | \hat{X} | \xi_m \rangle_{\text{app}}$. \tilde{X}_n is obtained by experimental data and $\langle \xi_m | \hat{X} | \xi_m \rangle_{\text{app}}$ is by theoretical calculation. We have considered that both of these two components contained some errors. This motivates us to classify the difference (4.2) in two components of errors according to the inequality,

$$\epsilon_{\text{op}}(n) \leq \epsilon_{\text{th}} + \epsilon_{\text{exp}}(n). \quad (4.3)$$

ϵ_{th} is the *theoretical error*, which is the error contained in $\langle \xi_m | \hat{X} | \xi_m \rangle_{\text{app}}$, and $\epsilon_{\text{exp}}(n)$ is the *experimental error*, which is the error contained in \tilde{X}_n . The reason why (4.3) is not an equation but an inequality is that the elements may cancel each other out.

In such an error evaluation, in many cases, after considering each factor, the square root of the sum of squares is taken by considering that each of these elements is uncorrelated. However, when there is a correlation between them, a larger error can be found. In the spirit of [36], we can evaluate the maximum possible error by using the triangle inequality. Of course, all factors need to be considered in order to evaluate the maximum value of the error that can really occur. Therefore, if an error that cannot occur in our analysis is found in an actual experiment, it indicates that the error factors that we do not consider are needed to explain the error in the experiment.

The theoretical error ϵ_{th} in (4.8) can also be decomposed into two distinct components under the inequality $\epsilon_{\text{th}} \leq \epsilon_{\text{mod}} + \epsilon_{\text{app}}$ with the *model error*,

$$\epsilon_{\text{mod}} := \left| \langle \xi_m | \hat{X} | \xi_m \rangle - X_\infty \right|, \quad (4.4)$$

and the *approximation error*,

$$\epsilon_{\text{app}} := \left| \langle \xi_m | \hat{X} | \xi_m \rangle_{\text{app}} - \langle \xi_m | \hat{X} | \xi_m \rangle \right|, \quad (4.5)$$

where $\langle \xi_m | \hat{X} | \xi_m \rangle$ is the theoretical expectation value calculated strictly. In our case, the experiment should be modeled by using the quantum mechanics so that we consider the expectation value should be written by the form $\langle \xi_m | \hat{X} | \xi_m \rangle$.

Analogously, to evaluate the experimental error, first we consider the ‘true’ values $\{x_i | i = 1, \dots, n\}$. These values would be obtained by carrying out the experiment with perfect accuracy. These values give us the ‘true’ average

$$X_n = \frac{1}{n} \sum_{i=1}^n x_i. \quad (4.6)$$

When the experiment is repeated infinitely many times, the true average X_n tends to the limit

$$X_\infty := \lim_{n \rightarrow \infty} X_n. \quad (4.7)$$

The limit (4.7) must be equal to the theoretical expectation value $\langle \xi_m | \hat{X} | \xi_m \rangle$ if our theoretical model is perfect. Based on the above discussions, it is appropriate to give the experimental error and theoretical error as follows.

$$\epsilon_{\text{exp}}(n) := |\tilde{X}_n - X_\infty|, \quad \epsilon_{\text{th}} := \left| \langle \xi_m | \hat{X} | \xi_m \rangle_{\text{app}} - X_\infty \right|. \quad (4.8)$$

The experimental error ϵ_{exp} in (4.8) is further decomposed into two distinct components according to the inequality

$$\epsilon_{\text{exp}}(n) \leq \epsilon_{\text{int}}(n) + \epsilon_{\text{st}}(n), \quad (4.9)$$

with the *statistical error*,

$$\epsilon_{\text{st}}(n) := |X_n - X_\infty|, \quad (4.10)$$

and the *intractable error*,

$$\epsilon_{\text{int}}(n) := |\tilde{X}_n - X_n|. \quad (4.11)$$

This intractable error is the difference between the true value and the actually obtained value.

Summing up these types of errors provide the upper bound for the difference (4.2)

$$\epsilon_{\text{op}}(n) \leq \epsilon_{\text{mod}} + \epsilon_{\text{app}} + \epsilon_{\text{st}}(n) + \epsilon_{\text{int}}(n). \quad (4.12)$$

In our discussion, we evaluate the appropriate size of each error with the method shown in the following section. It is more appropriate to say that what these estimates give is not an error but a kind of uncertainty caused by each effect. We use the sum of the estimated uncertainties for the total uncertainty Γ ,

$$\Gamma := \delta_{\text{mod}} + \delta_{\text{app}} + \delta_{\text{int}} + \delta_{\text{st}}. \quad (4.13)$$

δ_{int} , δ_{st} , δ_{mod} , and δ_{app} are the uncertainties corresponding to $\epsilon_{\text{int}}(n)$, $\epsilon_{\text{st}}(n)$, ϵ_{mod} , and ϵ_{app} , respectively.

We also evaluate the *relative uncertainty*, defined by the ratio of the shift in the actual experiment to the total uncertainty:

$$R := \frac{\Gamma}{\Delta_{\text{w}}(\theta)}. \quad (4.14)$$

$\Delta_{\text{w}}(\theta)$ is defined by (3.30). With this R , one may conclude that *the result of the measurement is valid if $R < 1$ and invalid otherwise*.

4.2 Calculation model

In this section, we give a method to evaluate every four uncertainties δ_{int} , δ_{st} , δ_{mod} , and δ_{app} explicitly. Since we have tried to apply a proposal [36] to the actual experiments, the following contents are based on the proposal.

4.2.1 Model uncertainty

To evaluate the model uncertainty δ_{mod} , we need to understand the model error defined by (4.4). If we cannot construct the theory correctly, the theoretical expectation value $\langle \xi_m | \hat{X} | \xi_m \rangle$ cannot be equal to X_∞ .

It may be considered that when we find this error, the model should be rejected, and it seems unnecessary to introduce this error. However, if there are many alternative theoretical models, this problem can exist as a kind of uncertainty.

Although this error is also included in the right-hand side of (4.12), considering this essence, it is appropriate for experimental science to treat this error as non-existent. Therefore, it is assumed that there is no such error in the following analysis.

4.2.2 Approximation uncertainty

In general, even if we choose a calculation model, it may not be possible to complete a calculation by this model. In many cases, we obtain the final numerical results using some approximations at some stages. Such approximations cause uncertainty. To evaluate this uncertainty, we need to compare two different values corresponding to the same expectation value $\langle \xi_m | \hat{X} | \xi_m \rangle$. In order to obtain these values, we calculate the expectation value $\langle \xi_m | \hat{X} | \xi_m \rangle$ with two different approximation methods.

The displacement using the WVA (3.33) is calculated with the first-order approximation of the parameter θ . This first-order approximation brings a simple result. This first-order approximation has been introduced in many papers as an equation to introduce the WVA. However, we can calculate (3.33) in another way. When we assume that the Gaussian function,

$$G(x) = \left(\frac{2\alpha}{\pi} \right)^{\frac{1}{4}} \exp(-\alpha x^2), \quad (4.15)$$

is applied to the wave function of the meter system $\chi(x)$, we can also calculate $\langle \xi_m | \hat{X} | \xi_m \rangle$, in which case it is possible to eliminate the first-order approximation.

We consider the difference of them as the approximation uncertainty,

$$\delta_{\text{app}} = \langle \xi_m | \hat{X} | \xi_m \rangle_{\text{G}} - \langle \xi_m | \hat{X} | \xi_m \rangle_{\text{1st}}. \quad (4.16)$$

$\langle \xi_m | \hat{X} | \xi_m \rangle_{\text{G}}$ is the expectation value calculated with using the Gaussian wave function (4.15) as $\langle x | \chi_{\text{pre}} \rangle$. Also, $\langle \xi_m | \hat{X} | \xi_m \rangle_{\text{1st}}$ is the expectation value calculated with using the first-order approximation of θ .

Comments on the validity of using the Gaussian as a wave function in experiments using laser beams

In optics, the paraxial approximation is well known as a very good approximation to calculate the electric field describing the laser beam [70, 71]. When this approximation is applied to the wave function of the electric field, we obtain some solutions, including the Gaussian function. (See appendix A) Also, using a quantum-classical correspondence, this classical electric field is considered as the wave function of the meter system. Combining them, we expect that the assumption using the Gaussian wave function (4.15) gives a good model for describing experiments we analyze. This means

$$\langle \xi_m | \hat{X} | \xi_m \rangle_{\text{G}} = \langle \xi_m | \hat{X} | \xi_m \rangle, \quad (4.17)$$

so that we do not distinguish them and we write them as $\langle \xi_m | \hat{X} | \xi_m \rangle$ unless it is necessary to make a distinction.

Evaluation in our analysis

The Gaussian wave function (4.15) for the position of the photon is not an eigenstate of the free Hamiltonian. Therefore, it is necessary to determine when the form should be applied. The present author compared the theoretical model of the time evolution of this optical system, that is, the discussion of Fourier optics [71] and the errors in the WVA [36], and considered that the wave function (4.15) should be modified and dealt with as follows.

In our analysis, we used the Gaussian wave function (4.15) for $\langle x|\chi_{pre}\rangle$. To derive the state $|\chi\rangle$,

$$\begin{aligned} |\chi\rangle &= e^{-\frac{i}{2k_0}z\hat{P}^2} e^{-i\frac{k_0}{2l_f}\hat{x}^2} |\chi_{pre}\rangle \\ &= e^{-\frac{i}{2k_0}z\hat{P}^2} e^{-i\frac{k_0}{2l_f}\hat{X}^2} \int dx G(x)|x\rangle \\ &= e^{-\frac{i}{2k_0}z\hat{P}^2} \int dx \left(\frac{2\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\left(\alpha+i\frac{k_0}{2l_f}\right)x^2} |x\rangle. \end{aligned} \quad (4.18)$$

To calculate (4.18), we insert the identity

$$\int dp |p\rangle \langle p|, \quad (4.19)$$

that is,

$$\begin{aligned} |\chi\rangle &= \int dp dx \left(\frac{2\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{i}{2k_0}zp^2} e^{-\left(\alpha+i\frac{k_0}{2l_f}\right)x^2} |p\rangle \langle p|x\rangle \\ &= \int dp \left(\frac{2\alpha}{\pi}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2\alpha+i\frac{k_0}{l_f}}} e^{-\left(i\frac{z}{2k_0}+\frac{1}{4\alpha+i\frac{2k_0}{l_f}}\right)p^2} |p\rangle, \end{aligned} \quad (4.20)$$

or its position base representation

$$\begin{aligned} |\chi\rangle &= \int dx \left(\frac{2\alpha}{\pi}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2\alpha+i\frac{k_0}{l_f}}} \frac{1}{\sqrt{2\left(i\frac{z}{2k_0}+\frac{1}{4\alpha+i\frac{2k_0}{l_f}}\right)}} e^{-\frac{1}{4\left(i\frac{z}{2k_0}+\frac{1}{4\alpha+i\frac{2k_0}{l_f}}\right)}x^2} |x\rangle \\ &= \int dx \left(\frac{2\alpha}{\pi}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2\alpha+i\frac{k_0}{l_f}}} \sqrt{\frac{k_0(2\alpha l_f + ik_0)}{k_0(l_f - z) + 2i\alpha z l_f}} e^{-\frac{k_0(2\alpha l_f + ik_0)}{2k_0(l_f - z) + 4i\alpha z l_f}x^2} |x\rangle \\ &= \int dx \left(\frac{2\alpha}{\pi}\right)^{\frac{1}{4}} \sqrt{\frac{k_0 l_f}{k_0(l_f - z) + 2i\alpha z l_f}} e^{-\frac{k_0(2\alpha l_f + ik_0)}{2k_0(l_f - z) + 4i\alpha z l_f}x^2} |x\rangle. \end{aligned} \quad (4.21)$$

Using (3.80), we can calculate the state $|\xi\rangle$ exactly. We will derive the results of applying \hat{X} and \hat{P} to \hat{Y} so that they can be applied to each experiment. In $\hat{Y} = \hat{X}$ case, except for x -independent constants, the state of the meter system after the postselection is

$$\langle \phi|\hat{U}(\theta)|\psi\rangle|\chi\rangle \propto \left[\cos(\theta\hat{X}) - iA_w \sin(\theta\hat{X})\right] \int dx e^{-\frac{k_0(l_f+2ik_0\alpha)}{4k_0\alpha(l_f-z)+2izl_f}x^2} |x\rangle, \quad (4.22)$$

$$= \int dx (\cos(\theta x) - iA_w \sin(\theta x)) e^{-\frac{k_0(l_f+2ik_0\alpha)}{4k_0\alpha(l_f-z)+2izl_f}x^2} |x\rangle =: |\zeta\rangle. \quad (4.23)$$

Also, in $\hat{Y} = \hat{P}$ case,

$$\langle \phi | \hat{U}(\theta) | \psi \rangle | \chi \rangle \propto \left[\cos(\theta \hat{P}) - i A_w \sin(\theta \hat{P}) \right] \int dp e^{-\frac{4k_0 \alpha (l_f - z) + 2izl_f}{4k_0 (l_f + 2ik_0 \alpha)} p^2} |p\rangle, \quad (4.24)$$

$$= \int dp (\cos(\theta p) - i A_w \sin(\theta p)) e^{-\frac{4k_0 \alpha (l_f - z) + 2izl_f}{4k_0 (l_f + 2ik_0 \alpha)} p^2} |p\rangle =: |\zeta_P\rangle. \quad (4.25)$$

Although these $|\zeta\rangle$ and $|\zeta_P\rangle$ are not normalized, they are useful in our calculation.

4.2.3 Statistical uncertainty

The statistical error $\epsilon_{\text{st}}(n)$ derives from the finiteness of the number of measurements performed in the experiment. The difference between the mean value of the results of the measurement of the pointer variable and the expectation value is known to be proportional to the standard deviation and inversely proportional to the square root of the number of measurement outcomes. For a quantitative evaluation of the deviation, which is indispensable for our definition of statistical uncertainty, we invoke *Chebychev's inequality* in its relevant form:

$$\Pr[\epsilon_{\text{st}}(n) \leq \delta_{\text{st}}] \geq T(\delta_{\text{st}}; n). \quad (4.26)$$

The inequality states that the probability distribution $\Pr[\epsilon_{\text{st}}(n) \leq \delta_{\text{st}}]$ of the deviation to be less than a bound $\delta_{\text{st}} > 0$ can be evaluated from below by the lower-bound function,

$$T(\delta_{\text{st}}; n) := \max \left[1 - \frac{(\Delta x)^2}{n \delta_{\text{st}}^2}, 0 \right], \quad (4.27)$$

with the variance of the distribution of the true values,

$$(\Delta x)^2 := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_i - X_\infty)^2. \quad (4.28)$$

To proceed, let $\eta \in [0, 1]$ be our assurance level of measurement, namely, the statistical uncertainty is assured to be smaller than a certain bound δ_{st} with probability η . From Chebychev's inequality (4.26), we can put $\eta = T(\delta_{\text{st}}; n)$, which may be used to determine the bound δ_{st} by solving the condition

$$\eta = 1 - \frac{(\Delta x)^2}{n \delta_{\text{st}}^2}, \quad (4.29)$$

unless n is too small or $(\Delta x)^2$ is too large for which $T(\delta_{\text{st}}; n)$ vanishes. We are thus left with the bound for the statistical uncertainty,

$$\epsilon_{\text{st}} = |X_n - X_\infty| \leq \sqrt{\frac{(\Delta x)^2}{n(1-\eta)}}. \quad (4.30)$$

We can use (4.30) as the statistical uncertainty δ_{st} . To obtain the statistical uncertainty, we need to give a reasonable assurance level η . Surely, it cannot be 1. As one index for assurance level, the 1σ interval of the normal distribution is often used. However, there are also research fields that consider this assurance level up to the 5σ interval. As such, the assurance level requirements vary from situation to situation and are not

absolute. In reference to [54], 0.95 is adopted to η in correspondence with the 2σ interval of the normal distribution.

Although we cannot know the variance $(\Delta x)^2$ given in (4.28) without performing the observations infinitely many times, we may instead use the theoretical variance evaluated for the meter state $|\xi_m\rangle$, namely,

$$(\Delta x)^2 = \langle \xi_m | \hat{X}^2 | \xi_m \rangle - \langle \xi_m | \hat{X} | \xi_m \rangle^2, \quad (4.31)$$

which holds true when our theoretical model is exact also in the variance.

Number fluctuation caused by postselection

When we consider experiments without postselection, the number of statistics is equal to the number of the repetitions of the experiments. However, this is not true in experiments involving the postselection. In order to know the number of statistics, we have to obtain the actual results. Also, we cannot expect it strictly by theoretical expectation in advance. Besides, there is a case that the experimenters make only the number of repetitions of the experiments public.

In order to deal with this problem, a method has been proposed [36]. This method takes cautionary mathematical points of the conditional probability into consideration. We have used this method in our analysis in order to consider the points properly.

To accommodate the possible fluctuation of output number n , we multiply both sides of (4.26) by the probability $P(n)$, and sum over n from zero to N to obtain

$$\Pr[\epsilon_{\text{st}} \leq \delta_{\text{st}}] \geq r(\delta_{\text{st}}; N, q) := \sum_{n=1}^N T(\delta_{\text{st}}; n) P(n). \quad (4.32)$$

At this point, we observe that, for our case of the probability distribution $P(n)$ being given by the binomial distribution

$$P(n) = \text{Bi}(n; N, q) := \frac{N!}{n!(N-n)!} q^n (1-q)^{N-n}, \quad (4.33)$$

where q is the postselection rate defined by (3.22). It is intuitively clear that the lower-bound function $r(\delta_{\text{st}}; N, q)$ defined in (4.32) is a monotonically increasing function with respect to all of its parameters δ_{st} , N and q (for a proof, see Appendix B.3). Employing a similar argument used for solving η in favor of δ_{st} in (4.29), thanks to the monotonicity of the function $r(\delta_{\text{st}}; N, q)$ we can solve

$$\eta = r(\delta_{\text{st}}; N, q), \quad (4.34)$$

to obtain its solution

$$\delta_{\text{st}}(\eta; N, q). \quad (4.35)$$

In our following analysis, we use (4.35) as the statistical uncertainty.

We can transform (4.32) into

$$\epsilon_{\text{st}} \leq \delta_{\text{st}}(\eta; N, q), \quad (4.36)$$

with using (4.35). This shows that on average the statistical error is guaranteed to be bounded from above with probability not less than η . Due to the monotonicity of $r(\delta_{\text{st}}; N, q)$, it is straightforward to see that $\delta_{\text{st}}(\eta; N, q)$ increases monotonically with respect to η , while it decreases monotonically with respect to N and q .

4.2.4 Intractable uncertainty

We define the intractable error as the difference between the actually obtained value in the experiment and the true value. In order to estimate this error and evaluate its corresponding uncertainty (intractable uncertainty) δ_{int} , we refer to an idea mentioned in the paper [36]. In this paper, the authors mentioned ‘*we can never have perfect knowledge of every technical imperfection in actual measurements, and consequently we may never know what the intended ‘true’ value is.*’

Technical imperfections can often be described as some physical effects. If we understand such physical effects as causes of uncertainty, uncertainty caused by technical imperfections may be understood as that associated with the selection of theoretical models that take into these physical effects. In this case, we can understand the uncertainty caused by technical imperfections as a kind of theoretical uncertainty, in particular, model uncertainty. However, it is extremely difficult to consider all such factors, and it is not always possible to correct the probability distribution. Rather, we do not consider unnecessarily such complicated things by presetting the ‘performance’ of the measuring instrument and the ‘applicability’ of the experimental method. Therefore, the meaning of the intractable error is the performance of such an experimental system.

When we deal with this uncertainty, it is inevitable to consider the following two problems. One is how to model describing the error. The other is how we evaluate the size in the actual system.

For the first problem, the paper [36] gives a model. This is called δ -uncertainty model. More specifically, to treat error explicitly, let $\{\tilde{x}_1, \dots, \tilde{x}_n\}$ be the n outcomes obtained by the measurements of the pointer variable X . Also, we assume that the following condition is satisfied between these outcomes and the true values x_i .

$$x_i \in [\tilde{x}_i - \delta_X, \tilde{x}_i + \delta_X] \text{ for } n = 1, \dots, n, \quad (4.37)$$

The parameter δ_X in (4.37) is a parameter representing the performance of the experimental system. We consider this δ_X as δ_{int} defined in (4.13). The error ϵ_{int} is evaluated according to (4.37) as follows.

$$\epsilon_{\text{int}} = |\tilde{X}_n - X_n| = \left| \frac{1}{n} \sum_{i=1}^n (\tilde{x}_i - x_i) \right| \leq \frac{1}{n} \sum_{i=1}^n |\tilde{x}_i - x_i| \leq \frac{1}{n} \sum_{i=1}^n \delta_{\text{int}} = \delta_{\text{int}}. \quad (4.38)$$

Therefore, it is assumed that δ_{int} given in (4.37) is applied to δ_{int} in (4.13). The inequality (4.38) is maximized when there is a strong correlation between the outcomes, and there is a systematic error that all deviate uniformly. We would like to revisit this kind of correlation issue in Section 4.6.

For the second problem, we have to determine an appropriate δ_{int} with fitting the conditions of the experiment we analyze consistently. Due to the ambiguity of the original idea, this error can include not only the performance of the measuring instrument for the pointer variable X but also some experimental technical problems. This point makes it difficult to determine δ_{int} . For the determination, we follow an idea from Morisawa [72]. In the idea, we may adopt the size of the error bars given by the experimenters as δ_{int} .

As we will see in Section 4.5, it is expected that the statistical uncertainty δ_{st} for which the quantitative evaluation method has already been given is smaller than the errors described in the experimental results. Besides, the experiments we analyzed were performed prior to the proposals of our approximation uncertainty and similar problems. Due to this, it is hard to expect that the error given by the experimenters include our approximation uncertainty. This means that the uncertainties given in the experiments

we analyzed are expected to be mainly corresponding to the intractable uncertainty in our classification. Due to this reason, the present author judged that Morisawa's idea is reasonable.

4.3 WVA in individual experiments

So far, we have explained the matters common to the WVA. In order to estimate the actual errors, we have to evaluate some values specific to individual experiments. In this section, we explain these points.

4.3.1 SHEL experiment

In the SHEL experiment, the experimenters wanted to measure the transverse deflection of the beam caused by the spin Hall effect. Therefore, the spin Hall effect can be understood as something that causes a transverse deflection depending on the spin of the photon. In order to describe the case, we can adopt the momentum operator \hat{P} to the operator \hat{Y} in (3.2). Then, the interaction $\hat{U}(\theta)$ becomes the operator describing the SHEL experiment and θ becomes the parameter we just want to know.

Using (3.18), (3.76), (3.78), (4.25), and some results in Appendix B.5, the transverse shift of the position of the beam is

$$\begin{aligned} \langle \xi_m | \hat{X} | \xi_m \rangle &= \langle \xi | \hat{X} | \xi \rangle + \frac{z_2}{k_0} \langle \xi | \hat{P} | \xi \rangle \\ &= \theta \frac{\text{Re}[A_w] + \text{Im}[A_w] \frac{\text{Im}[\alpha'_P]}{\text{Re}[\alpha'_P]} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}} + \frac{z_2}{k_0} \frac{\text{Im}[A_w] \frac{\theta}{2\text{Re}[\alpha'_P]} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}} \\ &= \theta \frac{\text{Im}[A_w] e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \left(\frac{\text{Im}[\alpha'_P]}{\text{Re}[\alpha'_P]} + \frac{z_2}{2k_0 \text{Re}[\alpha'_P]} \right)}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}} + \theta \frac{\text{Re}[A_w]}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}. \end{aligned} \quad (4.39)$$

In this experiment, the weak value is pure imaginary. Addition to this, the interaction is performed at the focal point of the lens. Using them, (4.39) become more simple form

$$\langle \xi_m | \hat{X} | \xi_m \rangle = \theta \left(\frac{\text{Im}[\alpha'_P]}{\text{Re}[\alpha'_P]} + \frac{z_2}{2k_0 \text{Re}[\alpha'_P]} \right) \frac{\text{Im}[A_w] e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}{1 + \frac{1}{2} (\text{Im}[A_w]^2 - 1) \left(1 - e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \right)}. \quad (4.40)$$

To evaluate the relative uncertainty R defined by (4.14), we can use

$$\Delta_w(\theta) = \langle \xi_m | \hat{X} | \xi_m \rangle. \quad (4.41)$$

Because the origin of the pointer variable is adjusted such that the parameter θ is zero, the shift is equal to the expectation value.

In our analysis, to evaluate the approximation uncertainty, we apply the first-order approximation of θ to (4.40), then we obtain

$$\langle \xi_m | \hat{X} | \xi_m \rangle_{1\text{st}} = \theta \left(\frac{\text{Im}[\alpha'_P]}{\text{Re}[\alpha'_P]} + \frac{z_2}{2k_0 \text{Re}[\alpha'_P]} \right) \text{Im}[A_w]. \quad (4.42)$$

To evaluate the statistical uncertainty, we need the expectation value of \hat{X}^2 . This is

$$\begin{aligned} \langle \xi_m | \hat{X}^2 | \xi_m \rangle &= \text{Re}[\alpha'_P] + \frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]} + \frac{z_2 \text{Im}[\alpha'_P]}{k_0 \text{Re}[\alpha'_P]} + \frac{z_2^2}{4k_0^2 \text{Re}[\alpha'_P]} \\ &+ \theta^2 \frac{\frac{1+\text{Im}[A_w]^2}{2} - \frac{1-\text{Im}[A_w]^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \left(\frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]^2} + \frac{z_2 \text{Im}[\alpha'_P]}{k_0 \text{Re}[\alpha'_P]^2} + \frac{z_2^2}{4k_0^2 \text{Re}[\alpha'_P]^2} \right)}{1 + \frac{1}{2} (\text{Im}[A_w]^2 - 1) \left(1 - e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \right)}. \end{aligned} \quad (4.43)$$

In this experiment, the following parameters are used.

$$\begin{aligned} z_1 &= 25 \text{ mm} \\ z_2 &= 125 \text{ mm} \\ k_0 &= 9.93 \text{ } \mu\text{m}^{-1} \end{aligned}$$

In the experimental paper [18], the information on the parameter α is given by

$$F = \frac{z_2}{2k_0 \text{Re}[\alpha'_P]} = 157 \pm 6.$$

We consider this F is strictly equal to 157 to obtain $\text{Re}[\alpha'_P]$.

4.3.2 USBD experiment

In the USBD experiment, what we want to detect is shift of the beam direction caused by the tilt of the mirror. In order to describe the case, we can adopt the position operator \hat{X} to the operator \hat{Y} in (3.2). Then, the interaction $\hat{U}(\theta)$ becomes the operator describing the USBD experiment and θ becomes the parameter we just want to know.

Using the distance from the lens to the mirror along the optical path is z_1 and the distance from the mirror to the detector is z_2 , we calculate the shift on the detector as follows:

$$\begin{aligned} \langle \xi_m | \hat{X} | \xi_m \rangle &= \langle \xi | \hat{X} | \xi \rangle + \frac{z_2}{k_0} \langle \xi | \hat{P} | \xi \rangle \\ &= \frac{\text{Im}[A_w] \frac{\theta}{2\text{Re}[\alpha']} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}} + \frac{z_2}{k_0} \frac{-\theta \text{Re}[A_w] - \theta \text{Im}[A_w] \frac{\text{Im}[\alpha']}{\text{Re}[\alpha']} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}} \\ &= \frac{\text{Im}[A_w] \frac{\theta}{\text{Re}[\alpha']} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(\frac{1}{2} - \frac{z_2}{k_0} \text{Im}[\alpha'] \right) - \theta \frac{z_2}{k_0} \text{Re}[A_w]}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}} \\ &= \frac{\text{Im}[A_w] \frac{\theta}{\text{Re}[\alpha']} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(\frac{1}{2} - \frac{z_2}{k_0} \text{Im}[\alpha'] \right) - \theta \frac{z_2}{k_0} \text{Re}[A_w]}{1 + \frac{-1+|A_w|^2}{2} (1 - e^{-\frac{\theta^2}{2\text{Re}[\alpha']}})}, \end{aligned} \quad (4.44)$$

where α' is given by (B.25). In order to evaluate the relative uncertainty R defined by (4.14), we can also use the relation (4.41) due to the same reason in the case of the SHELL experiment. Also, the expectation value with the first order approximation of θ is

$$\langle \xi_m | \hat{X} | \xi_m \rangle_{1\text{st}} = \theta \frac{\frac{1}{2} - \frac{z_2}{k_0} \text{Im}[\alpha']}{\text{Re}[\alpha']} \text{Im}[A_w]. \quad (4.45)$$

In our discussion, we need one more information on the variance of the experimental results in order to estimate the statistical error. To obtain this variance, we calculate the expectation value of \hat{X}^2

$$\begin{aligned} \langle \xi_m | \hat{X}^2 | \xi_m \rangle = & \frac{1}{4\text{Re}[\alpha']} - \frac{z_2}{k_0} + \frac{z_2^2}{k_0^2} \left(\text{Re}[\alpha'] + \frac{\text{Im}[\alpha']^2}{\text{Re}[\alpha']} + \theta^2 \right) \\ & - \frac{\theta^2}{4\text{Re}[\alpha']^2} \frac{\frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(1 - \frac{4z_1}{k_0} \text{Im}[\alpha'] + \frac{4z_2^2}{k_0^2} \text{Im}[\alpha']^2 \right)}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}. \end{aligned} \quad (4.46)$$

To obtain the actual values, we have used the following parameters for analyzing this experiment.

$$\begin{aligned} z_1 &= 48 \text{ cm} \\ z_2 &= 114 \text{ cm} \\ k_0 &= 8.06 \mu\text{m}^{-1} \\ \alpha &= 1.22 \times 10^{-6} (\mu\text{m})^{-2} \end{aligned}$$

At the end of this section, we explain some points on the beamwidth. The experimenters are also interested in the relation between the WVA and the beamwidth on the measuring instrument due to the result (3.34). Because of this motivation, the beamwidth at the detector σ_d has been chosen from about 300 μm to 1300 μm (see Figure 4.1) [19]. In order to change σ_d , the experimenter has not fixed the focal length l_f .

This l_f is required to calculate α' in (4.44) and (4.45). However, the experimenters give the values of σ_d instead of l_f . Although we can convert σ_d to l_f , it is a bit intricate. In order to calculate α' , we can use a simple formula

$$\sigma = \frac{z_1 a + \sigma_d z_2}{z_1 + z_2}. \quad (4.47)$$

This σ is the beamwidth at the tiltable mirror in Figure 4.1. Combining with a relation

$$\sigma = \frac{1}{\sqrt{2\text{Re}[\alpha']}}, \quad (4.48)$$

we can calculate $\text{Re}[\alpha']$ without using l_f . The parameter a in Figure 4.1 is

$$a = \frac{1}{\sqrt{2\alpha}} = 640 \mu\text{m}. \quad (4.49)$$

4.4 Parameter settings for simulation of each experiment

In order to calculate the uncertainties, we calculate the values $\langle \xi_m | \hat{X} | \xi_m \rangle$, ϵ_{app} , $\delta_{\text{st}}(\eta; N, q)$ and fixed value δ_{int} defined in (4.37). Although our formulas to calculate them have a number of free parameters, these parameters are clearly stated in the experimental papers except two parameters N and δ_{int} . In order to determine these two parameters in the optical experiments, such as the SHEL and USBD experiments, we need some interpretations of them.

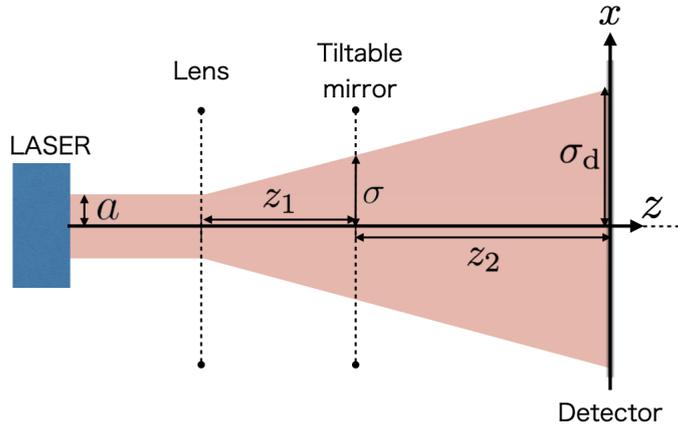


Figure 4.1: Expansion of the width of the Gaussian state of the meter by the lens and propagation. The original width a at the laser is expanded to σ at the tiltable mirror and further to σ_d at the detector. Here we have made the entire path of the photons into a straight line so that the expansion of the width along the passage is readily seen.

4.4.1 The number of statistics N

N is the number of statistics. In experiments using a laser beam that continuously emits photons, it cannot be uniquely determined for some reasons.

One is that it is not clear what we detect with our detector. In quantum-classical correspondence shown in Appendix A.1, we find the classical electric field of the beam can be considered as the wave function. However, if we consider that the detector is used for measuring the classical electric field, we cannot model the measurement as a quantum measurement precisely. More precisely, it is doubtful that each output of the detector is corresponding to the position for each photon. Therefore, the number of particles is not equal to the number of statics in general.

In practical meaning, we cannot ignore the *efficiency*. When we know the efficiencies in the actual systems, we should be able to estimate the actual number of statistics from the number of photons emitted from the light source. However, we cannot know these efficiencies in the actual experiments.

The number of photons can be estimated from the wavelength, the intensity of the laser beam, and exposure time. They have been described in the papers of these experiments. It is extremely large, exceeding 10^{13} [54]. Considering the burden on the computer and the difference in laser beam intensity between the SHEL experiment (10mW [18]) and the USBD experiment (3.2mW [19]), we adopt 10^9 as the value of N in the SHEL experiment. Also, we adopt 10^8 as the value of N in the USBD experiment.

4.4.2 The intractable uncertainty δ_{int}

As we mentioned in Subsection 4.2.4, we assign the error mentioned in the experimental paper to the intractable uncertainty δ_{int} .

SHEL experiment To determine the value of δ_{int} , we refer to Ref. [18]. It is stated that the unwanted shifts caused by rotating the polarizer are about the order of $10\mu\text{m}$, which appears to provide a dominant factor to determine the intractable uncertainty in

our discussion. From this we choose our parameter,

$$\delta_{\text{int}} = 10 \text{ } \mu\text{m}. \quad (4.50)$$

USBD experiment We used

$$\delta_{\text{int}} = 30 \text{ } \mu\text{m}, \quad (4.51)$$

in our analysis of the USBD experiment. This is the maximal size of the error bar attached to the measurement results shown in Fig.2 of [19].

4.5 Results of our analysis

We now examine, based on our theoretical framework, the two preceding experiments, the SHEL experiment by Hosten and Kwiat and the USBD experiment by Dixon *et al.*

We have calculated the theoretically expected shifts of the pointer variables and measurement uncertainties. Also, we have evaluated them together with the relative uncertainties described above.

In the SHEL and USBD experiments, we change other variables as well as weak values and verify their response. In the SHEL experiment, θ is changed by the angle of incidence of the beam on the prism, and in the USBD experiment, the beamwidth is manipulated by adjusting the focal length of the lens inserted into the beam. Therefore, we also have examined the behavior of the uncertainties when these variables are changed.

4.5.1 The SHEL experiment

In this experiment, the coupling θ depends on the setting of the variable angle prism, namely, the incident angle of the beam. The main result of the paper [18] is the confirmation of the dependence predicted theoretically when the spin Hall effect exists. We check the validity of this experiment in view of two points, the choice of the weak value (or postselection) and the range of the coupling employed.

Weak Value dependence In the experiment, the postselection is made so that $\text{Im}[A_w]$ is fixed to the value 57.3 ± 0.7 for the angle of the prism smaller than 56° , or to the value 31.8 ± 0.2 for the angle of the prism larger than 56° . In both cases, the ratio R is found well below 1, indicating that the result of the measurement is valid (see Figure 4.2).

At the selected weak value, the intractable uncertainty and approximation uncertainty are almost the same. The statistical uncertainty is actually expected to be smaller than this result because the number of trials N is significantly reduced in our simulation. Since the statistical uncertainty is not remarkable even after the number of statistics is significantly reduced in this way, it is expected that the statistical error will not be remarkable regardless of the selection of the weak value.

It is noteworthy that the value $\text{Im}[A_w] = 57.3 \pm 0.7$ is close to the optimal point 64.2 of the ratio R .

Coupling Dependence. In the experiment, the coupling constant is varied in the region $2 < \theta < 65$. We analyze the variation of the shift and uncertainties (Figure 4.3; left) and the relative uncertainties (Figure 4.3; right) in this region for θ at the fixed value of the weak value $\text{Im}[A_w] = 57$. From this, it is observed that the entire region of θ is

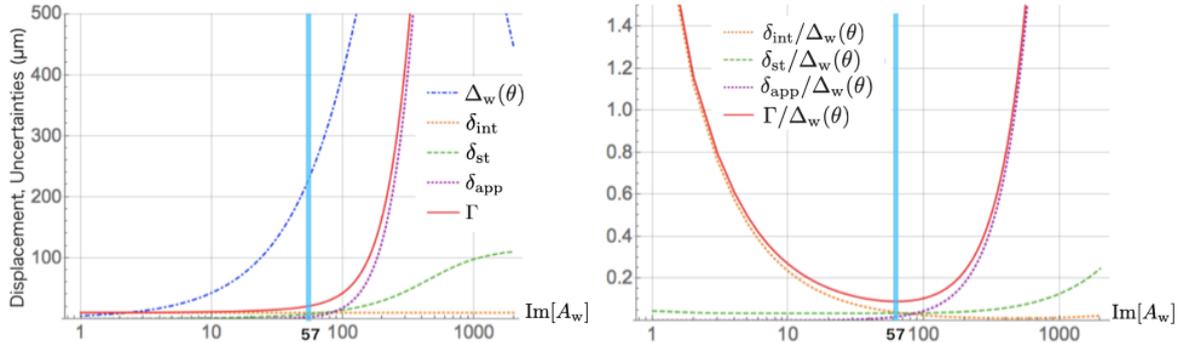


Figure 4.2: Shift and various uncertainties (left), the relative uncertainty $R = \Gamma/\Delta_w$ and its components (right) as functions of $\text{Im}[A_w]$, obtained at $g = 27\text{nm}$ for which the angle of the prism smaller than 56° . In this case, the actual value used in the experiment is $\text{Im}[A_w] = 57.3 \pm 0.7$ shown by the narrow strips colored in cyan.

in the safe zone of R of less than 1, and that it covers the optimal value of θ around 40. We thus find that the measurement as a whole is not just valid as precision measurement but is almost an optimal one, even though the measurements with values near $\theta = 2$ are almost at the lower limit according to our criterion.

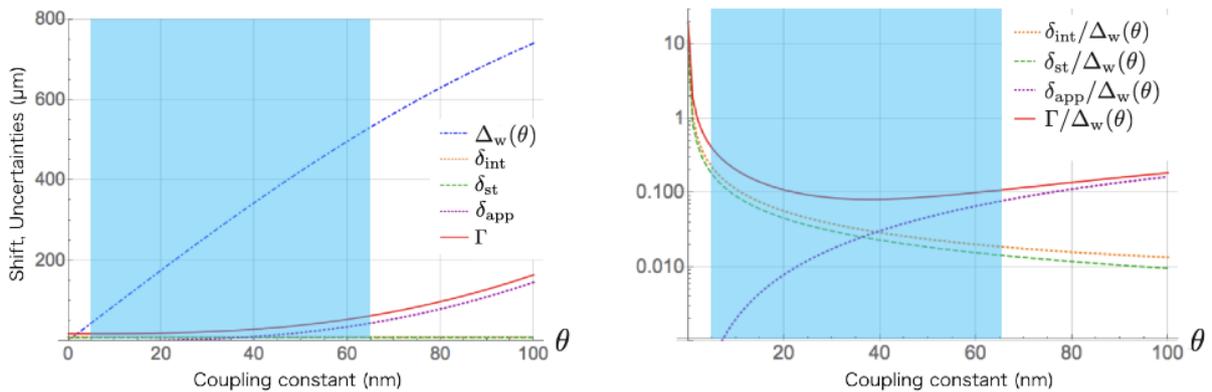


Figure 4.3: Shift and various uncertainties (left), the relative uncertainty $R = \Gamma/\Delta_w$ and its components (right) as functions of the coupling constant g , obtained at $\text{Im}[A_w] = 57$. The zones colored in cyan indicate the range of θ used in the experiment. The relative uncertainty R attains the minimum at around the value $\theta = 40$.

4.5.2 The USBD experiment

The width of the beam has been prepared to be adjustable in this experiment. Based on this, we have analyzed the shift and error changes when the beamwidth has been changed as well.

Weak Value Dependence In the experiment, the phase φ given in (3.92) was chosen so that the imaginary part of the weak value $\text{Im}[A_w]$ in (3.92) becomes 6.57, 9.93, 15.9. In Figure 4.4, the zone colored in cyan covers the three values of $\text{Im}[A_w]$ used in the experiment. Viewed from the shift $\Delta_w(k)$ and the relative uncertainty R , we immediately observe that, according to our criterion mentioned when we defined R before, the three values $\text{Im}[A_w]$ used in the experiment are in the safe region where the measurement is valid.

In fact, it is amusing to observe that the ratio R has its minima at around $\text{Im}[A_w] \approx 15$ and that the largest value $\text{Im}[A_w] = 15.9$ used in the experiment is found to be close to this optimal point.

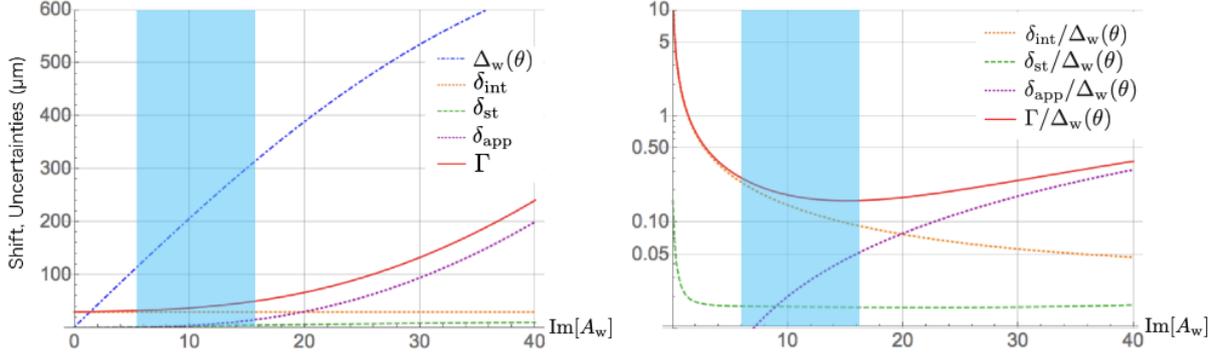


Figure 4.4: Shift/uncertainties (left) and relative uncertainties (right) as functions of $\text{Im}[A_w]$ obtained at $\sigma_d = 750\mu\text{m}$. The zones colored in cyan cover the actual values of $\text{Im}[A_w]$ used in the experiment.

Beamwidth Dependence. In the experiment, in addition to the value $\sigma_d = 750\mu\text{m}$, various other values in the region $300 < \sigma_d < 1300$ are also adopted for the beamwidth. We analyze the variation of the shift $\Delta_w(k)$ and the relative uncertainty R when we vary the width σ_d at the fixed value of the weak value $\text{Im}[A_w] = 10$ (see Figure 4.5). As before, for both the shift $\Delta_w(k)$ and the relative uncertainty R , the operational difference is determined primarily by the intractable uncertainty. Our analysis shows that using larger beamwidths can improve the relative uncertainty R if the technical problems can be removed.

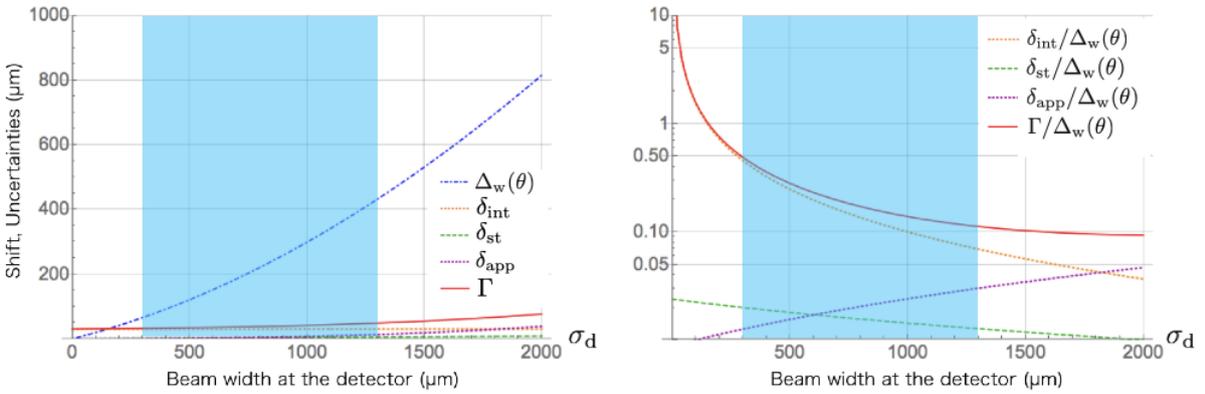


Figure 4.5: Shift and various uncertainties (left), the relative uncertainty $R = \Gamma/\Delta_w$ and its components (right) as functions of σ_d obtained at $\text{Im}[A_w] = 10$. The zones colored in cyan cover the actual values of σ_d used in the experiment.

4.6 Relevance to other studies

S. Pang and his collaborators have shown that the error caused by the decoherence of the meter system can be improved by postselection [37]. They evaluate an effect of some external disturbance (actually, decoherence on the meter system was selected and

analyzed) on the probability distribution and estimate the error as the linear response of the estimator to the parameters of such external disturbance. Comparing the case using the WVA with the case without the WVA, the WVA method can improve the performance of the parameter estimation.

In this analysis, they focus on the maximum likelihood estimation (MLE) method, which is often used to estimate parameters by comparing experimental results with parameter-dependent probability distributions. The maximum likelihood method is well used for analysis in many experiments. This method and our method are complementary. MLE analyzes the nature of the distribution very well and makes good use of the information obtained in the experiment.

In our research, it is suggested that the WVA method is superior to the performance limit of the measuring instrument used as a meter system for the measurement of the physical quantity of position. Other examples of research include cases where classical fluctuations with respect to the meter system are mixed [34], and cases where the meter system has decoherence [37]. Taken together, it is shown that the WVA can be superior, especially in certain environments where high performance cannot be required for the meter system.

A paper of Sinclair *et al.* [73] is also notable. In their paper, they consider the noise similar to our intractable error and they find the importance of the correlation of the noise in view of parameter estimation using the MLE method. This is also seen in our calculation because if there is no correlation between the individual errors of the output, number dependence also appears in this type of error. In fact, the expected intractable error is

$$\epsilon_{\text{int-nocorrelated}} := \frac{1}{n} \sqrt{\sum_{i=1}^n (\tilde{x}_i - x_i)^2} \leq \frac{1}{n} \sqrt{n \delta_{\text{int}}^2} = \frac{\delta_{\text{int}}}{\sqrt{n}}, \quad (4.52)$$

where n is the number of the actual measurement. In the WVA, this number is expected to be $N \times q$, where N is the number of trials and q is the postselection probability given by (3.22). Due to this number dependency, the behavior of the intractable error is expected to deteriorate when the weak value is taken as a variable becomes roughly proportional to the behavior of the statistical error, and as a result, the behavior of the error in the region where the weak value is large. It is expected to improve the error in the region where the weak value is small. If the intractable error that occurs for each measurement is uncorrelated (see Appendix B.4 for the reason), the error of the mean value is expressed by (4.52).

It is notable that some papers show the merit of the WVA with comparing the cases with and without the postselection (see *e.g.* [30, 34, 37, 66]). Each study has used its specific assumption. To maximize the merit, we should consider both the positive and negative sides of the WVA with a certain comparison with the experiments without the postselection.

4.7 Summary and Discussions of this chapter

Our analysis supports the reliability and validity of the results of both SHEL and USBD experiments. It can be said that this was the result of strengthening the effectiveness of the WVA. In particular, in the case of optical systems, it is shown that the statistical error is sufficiently small even if the number of particles is reduced by five orders of magnitude. This shows that our simple statistical error is not an essential problem in

actual experimental systems and that the statistical weakness of the WVA shown by Tanaka-Yamamoto [30] or Knee [31] is not essential in the actual optical systems. On the other hand, the following two problems become apparent.

First, although our method evaluates the maximum possible error as long as we can assume that all the error factors are included, the systematic error seen in the USBD experiment is much larger than that expected in our analysis. It seems that the systematic error is not a universal problem because it arises only for one of the cases of the weak values selected in the USBD experiment. However, it became more or less clear that the discussions considered by our analysis method did not exhaust the cause of the error in the actual experiment.

Second, there are some differences between the assumptions in a proposal [36] and the actual experiments. For example, the number of photons we evaluated in Section 4.4 ($\sim 10^{13}$) is too large to use in our computers. Of course, by recording only the number of particles captured for each detector without distinguishing each particle, the resources used can be significantly reduced. However, a new problem may occur related to the accuracy of the model used, which cannot be treated by their method without major changes.

In that sense, our method may be more relevant to systems with a smaller number of particles rather than those with a large number of particles such as optical systems.

Chapter 5

An application of WVA to high energy experiments

Most of discussions of the WVA has been based on the selection of the target system as a two-level system and the meter system as the position or momentum.

Since the interference length is expected to be extremely short in high energy physics, the WVA cannot be expected at all in the context of shifting the position. However, time coherence has been well verified and commonly used in high energy experiments. For example, neutrino oscillations and Bell's inequality verification experiments using mesons [53] also utilize changes in quantum states over time. Therefore, if weak values appear during time evolution, it is expected that the discussions using the weak value can be brought into high-energy experiments. This observation leads us to consider the possibility of weak values appearing during time evolution to find the possibility of the WVA in high-energy experiments.

We first explain that the weak value can be obtained from the particle decay in Section 5.1. We have tried to apply the WVA to the B meson. In Section 5.2, we explain the dynamics of B meson and obtain some results for our analysis. It is important that whether these results can work as the WVA. In our analysis, we can see the imaginary part of the weak value of the modified Hamiltonian in the prolongation of the effective lifetime. In Section 5.3, we have checked that the prolongation of the effective lifetime. We obtain these results by assuming that our postselection can be performed. However, the method is not constructed. In Section 5.4, we propose an idea and have checked what we need to perform our postselection. In Sections 5.5, 5.6, and 5.7, we consider the actual experimental conditions and estimate the effect of the WVA. Finally, we comment on some points in Section 5.8.

5.1 Weak Value in Particle Decay

Before we begin to discuss the application of the weak measurement to B meson systems, we describe the basic reason why the weak value arises in analyzing the particle decay in general.

In the context of particle decay, one may expect that the WVA can be used to prolong effectively the decay time of the initial particle by considering the Hamiltonian operator \hat{H} for the observable \hat{A} . That this is indeed the case can be seen as follows.

Let $|i\rangle$ be the preselected state of the initial particle and $|f\rangle$ be the postselected state of the decayed particle(s). The probability of transition between them during the time

period t will then be described by

$$P(t | i \rightarrow f) = |\langle f | e^{-it\hat{H}} | i \rangle|^2. \quad (5.1)$$

Suppose that the Hamiltonian \hat{H} is bounded as is usually the case for considering transitions among a finite set of energy levels. Noting that \hat{H} is not Hermitian (or self-adjoint) to take account of the decaying processes, we may define the ‘normalized’ operator of time development by

$$\hat{A} = \frac{2}{\Delta m} \left\{ \hat{H} - \left(m - \frac{i}{2}\Gamma \right) \right\}, \quad (5.2)$$

where m and Γ are the averages of the real part and the imaginary part of the eigenvalues of \hat{H} , respectively, and Δm is the range of the real part, *i.e.*, the difference between the largest and the smallest eigenvalues in the real part. In particular, if there are only two states which enter in the transition, then the operator \hat{A} in (5.2) may admit the form of the familiar Pauli matrices, especially $\hat{A} = \sigma_3$ in the energy diagonal basis.

Substituting \hat{H} with the normalized one \hat{A} in the transition amplitude, we find

$$\langle f | e^{-it\hat{H}} | i \rangle = e^{-\frac{1}{2}\Gamma t} e^{-imt} \langle f | e^{-ig\hat{A}} | i \rangle, \quad (5.3)$$

where

$$g = \frac{1}{2}\Delta m t = \frac{1}{2} \left(\frac{\Delta m}{\Gamma} \right) \Gamma t, \quad (5.4)$$

is the effective coupling constant. Now, from (5.3) we realize that the effective range of the time period t may be restricted by the order $O(\Gamma t) = 1$. Hence if $O(\Delta m/\Gamma) \ll 1$, then g can be regarded as a small parameter $O(g) \ll 1$. This allows us to employ the linear approximation,

$$\begin{aligned} \langle f | e^{-ig\hat{A}} | i \rangle &\simeq \langle f | \hat{1} - ig\hat{A} | i \rangle \\ &= \langle f | i \rangle (1 - igA_w) \simeq \langle f | i \rangle e^{-igA_w}, \end{aligned} \quad (5.5)$$

where A_w is the weak value (3.21) of the normalized operator \hat{A} in (5.2). Plugging this expression into (5.1), we obtain

$$P(t | i \rightarrow f) \simeq e^{-\Gamma t} |\langle f | i \rangle|^2 e^{2g\text{Im}[A_w]}, \quad (5.6)$$

$$= e^{-(\Gamma - \Delta m \text{Im}[A_w])t} |\langle f | i \rangle|^2. \quad (5.7)$$

This shows that, as long as the linear approximation is valid, the probability distribution of the decay is indeed affected by the weak value A_w (in the imaginary part), which can be freely altered by the combination of the preselection and the postselection we perform. Also, at this time, the effective collapse width is

$$\Gamma' = \Gamma - \Delta m \text{Im}[A_w], \quad (5.8)$$

and this means that the effective lifetime τ_{eff} is

$$\begin{aligned} \tau_{\text{eff}} &= \frac{1}{\Gamma'} = \frac{1}{\Gamma - \Delta m \text{Im}[A_w]} \\ &\simeq \frac{1}{\Gamma} \left(1 + \frac{\Delta m}{\Gamma} \text{Im}[A_w] + O((\Delta m/\Gamma)^2) \right). \end{aligned} \quad (5.9)$$

The extension of the effective lifetime is proportional to the imaginary part of the weak value \hat{A} in the first order of $\Delta m/\Gamma$.

Even if the coupling constant g cannot be regarded small, the system may allow one to evaluate the transition amplitude in full order to see if the selections affect the outcome. Later, we shall see explicitly that the decay of the B meson falls into that particular case.

5.1.1 Meter system like understanding of this type amplification

Although the time is not an observable in quantum mechanics but a dynamical parameter, this discussion can be understood similar to weak measurement using a kind of the meter system.

The state is described as follows by time evolution.

$$|\psi(t)\rangle = e^{-it\hat{H}} |\psi\rangle, \quad (5.10)$$

where \hat{H} is a Hamiltonian and $|\psi\rangle$ is an initial state of the system. In the discussion here, we consider that the particles are unstable particles, so the eigenvalues of the Hamiltonian have imaginary parts, but for the sake of simplicity, we assume that the imaginary parts are the same constant $-\Gamma/2$. Then, the Hamiltonian becomes normal (diagonalizable) operator,

$$\hat{H} = \hat{M} - i\frac{\Gamma}{2}. \quad (5.11)$$

\hat{M} is a Hermitian operator, and we call it the mass operator. The state (5.10) is

$$\begin{aligned} \langle\phi|\psi(t)\rangle &= \langle\phi|e^{-\frac{\Gamma}{2}t}e^{-it\hat{M}}|\psi\rangle \\ &= \sum_i c_i d_i^* e^{-\frac{\Gamma}{2}t} e^{-itm_i} |m_i\rangle, \end{aligned} \quad (5.12)$$

where $c_i = \langle m_i|\psi\rangle$, and $d_i = \langle m_i|\phi\rangle$. m_i and $|m_i\rangle$ are eigenvalues and eigenvectors of the mass operator satisfying

$$\hat{M}|m_i\rangle = m_i|m_i\rangle, \quad (5.13)$$

for each i . Here, we define a function of time $\psi_i(t)$ by

$$\psi_i(t) := \begin{cases} e^{-\frac{\Gamma}{2}t} e^{-itm_i}, & t \geq 0, \\ 0, & t < 0, \end{cases} \quad (5.14)$$

then the state (5.12) can be written as

$$\langle\phi|\psi(t)\rangle = \sum_i c_i d_i^* \psi_i(t). \quad (5.15)$$

Considering the Fourier transform of $\psi_i(t)$, we can obtain a function

$$\begin{aligned} \tilde{\psi}_i(\omega) &:= \int_{-\infty}^{\infty} \frac{dt}{\sqrt{2\pi}} \psi_i(t) e^{i\omega t} \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{\Gamma}{2} - i(\omega - m_i)} \\ &= \tilde{\psi}(\omega - m_i), \end{aligned} \quad (5.16)$$

where

$$\tilde{\psi}(\omega) = \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{\Gamma}{2} - i\omega}. \quad (5.17)$$

This ω is corresponding to the *invariant mass*. For those who measure the invariant mass for this system, such a state can be regarded as a superposition of the ‘*eigenstates*’ for the invariant mass. That is, the state is

$$|\xi\rangle = \int d\omega \sum_i c_i d_i^* \tilde{\psi}(\omega - m_i) |\omega\rangle, \quad (5.18)$$

where $|\omega\rangle$ is the ‘eigenstate’ of the invariant mass.

This state (5.18) is similar to the state (3.19) and this means that the invariant mass brings a view point like using the meter system.

5.2 Dynamics of B meson and Time distributions

In this section, we present our theoretical basis for applying the weak measurement for the detection of CP violation in the B meson system. In our case, the Hamiltonian \hat{H} takes the role of the observable \hat{A} , and we shall see that the imaginary part of the weak value H_w is related to the effective lifetime of the B meson. The aim of the weak measurement is then to find a source of CP violation in the conditional time distribution under the condition that we have B^0 at the initial time and B_{decay} at the time of the decay.

5.2.1 Hamiltonian

The state $|\psi\rangle$ which describes a neutral B meson is in general a superposition of the flavor eigenstates $|B^0\rangle$ and $|\bar{B}^0\rangle$ which are the CP conjugate to each other, that is, $|\bar{B}^0\rangle = \mathcal{CP}|B^0\rangle$ and $|B^0\rangle = \mathcal{CP}|\bar{B}^0\rangle$ under the CP transformation \mathcal{CP} for which $(\mathcal{CP})^2 = \hat{1}$.

Since mesons are unstable particles, the time evolution is phenomenologically described by a non-Hermitian Hamiltonian \hat{H} . The eigenstates of the Hamiltonian are given by

$$\hat{H}|B_L\rangle = \left(m_L - \frac{i}{2}\Gamma_L\right)|B_L\rangle, \quad (5.19)$$

$$\hat{H}|B_H\rangle = \left(m_H - \frac{i}{2}\Gamma_H\right)|B_H\rangle, \quad (5.20)$$

where $m_L, m_H, \Gamma_L, \Gamma_H$ are real positive numbers satisfying

$$m_L < m_H. \quad (5.21)$$

These eigenstates $|B_L\rangle$ and $|B_H\rangle$ are given by

$$|B_L\rangle = p|B^0\rangle + q|\bar{B}^0\rangle, \quad (5.22)$$

$$|B_H\rangle = p|B^0\rangle - q|\bar{B}^0\rangle, \quad (5.23)$$

where p, q are complex numbers satisfying $|p|^2 + |q|^2 = 1$. This is derived for the Hamiltonian by assuming CPT invariance. (see Section B.2.) It should be noted that these eigenstates are not orthogonal, so the Hamiltonian \hat{H} cannot be diagonalized.

Using (5.22) and (5.23), we find

$$|B^0\rangle = \frac{1}{2p}(|B_L\rangle + |B_H\rangle), \quad (5.24)$$

$$|\bar{B}^0\rangle = \frac{1}{2q}(|B_L\rangle - |B_H\rangle). \quad (5.25)$$

In high energy experiment, we decide the initial state whether the flavor of a particle is B^0 or \bar{B}^0 through a process called the flavor tagging [74, 75]. In general, the problem of the vulnerability to noise contamination in the preselection described in section 3.4.1 may be ignorable because the misidentification is not unavoidable for several reasons [75]. However, this study is to estimate the effect of applying the WVA to high energy systems. Therefore, these imperfections of preselection and postselection procedures should be considered as simply as possible. That is, they are considered negligible.

In this dissertation, the time at which this tagging is performed is referred to as ‘initial’, and the estimated time at which the particle decay is denoted by Δt . When the meson identified as B^0 at the initial time, it evolves into the state,

$$|B^0(\Delta t)\rangle = e^{-i\Delta t\hat{H}}|B^0\rangle, \quad (5.26)$$

immediately before the time of the decay. Similarly, when the particle was identified as \bar{B}^0 , we have

$$|\bar{B}^0(\Delta t)\rangle = e^{-i\Delta t\hat{H}}|\bar{B}^0\rangle. \quad (5.27)$$

For our later convenience, we note here that, as long as the dynamical development is concerned, the evolution (5.27) of the CP conjugate \bar{B}^0 is obtained immediately from (5.26) by interchanging $p \leftrightarrow q$. This can be readily confirmed from (5.19) and (5.20) by expressing the Hamiltonian in terms of the parameters p, q in the orthogonal basis of the flavor eigenstates $|B^0\rangle$ and $|\bar{B}^0\rangle$ as

$$\begin{aligned} \hat{H}(p, q) = & (m - \frac{i}{2}\Gamma) (|B^0\rangle\langle B^0| + |\bar{B}^0\rangle\langle \bar{B}^0|) \\ & - (\Delta m - \frac{i}{2}\Delta\Gamma) \left(\frac{p}{2q}|B^0\rangle\langle \bar{B}^0| + \frac{q}{2p}|\bar{B}^0\rangle\langle B^0| \right), \end{aligned} \quad (5.28)$$

using

$$m = \frac{m_L + m_H}{2}, \quad \Gamma = \frac{\Gamma_L + \Gamma_H}{2}, \quad (5.29)$$

and

$$\Delta m = m_H - m_L, \quad \Delta\Gamma = \Gamma_H - \Gamma_L. \quad (5.30)$$

One then immediately observes that the CP conjugation of the Hamiltonian results in the interchange of the parameters, $\mathcal{CP}\hat{H}(p, q)\mathcal{CP} = \hat{H}(q, p)$. It then follows that, given the transition amplitude $T(p, q) = \langle\alpha|e^{-i\Delta t\hat{H}(p, q)}|\beta\rangle$, the corresponding amplitude of the CP conjugate has the same property,

$$\begin{aligned} \bar{T}(p, q) &= \langle\bar{\alpha}|e^{-i\Delta t\hat{H}(p, q)}|\bar{\beta}\rangle = \langle\alpha|\mathcal{CP}e^{-i\Delta t\hat{H}(p, q)}\mathcal{CP}|\beta\rangle, \\ &= \langle\alpha|e^{-i\Delta t\mathcal{CP}\hat{H}(p, q)\mathcal{CP}}|\beta\rangle = T(q, p), \end{aligned} \quad (5.31)$$

as claimed.

5.2.2 Time distribution without postselection

Due to the instability of the particle, the norm of the particle's state is not always unity, as can be confirmed explicitly by

$$\begin{aligned}
& \langle B^0(\Delta t) | B^0(\Delta t) \rangle \quad (= \langle B^0 | e^{i\Delta t \hat{H}^\dagger} e^{-i\Delta t \hat{H}} | B^0 \rangle), \\
&= \frac{1}{4|p|^2} \left(e^{im_L \Delta t - \frac{\Gamma_L}{2} \Delta t} \langle B_L | + e^{im_H \Delta t - \frac{\Gamma_H}{2} \Delta t} \langle B_H | \right) \\
&\quad \cdot \left(e^{-im_L \Delta t - \frac{\Gamma_L}{2} \Delta t} | B_L \rangle + e^{-im_H \Delta t - \frac{\Gamma_H}{2} \Delta t} | B_H \rangle \right), \\
&= \frac{1}{4|p|^2} \left(e^{-\Gamma_L \Delta t} + e^{-\Gamma_H \Delta t} \right) + \frac{e^{-\Gamma \Delta t}}{4|p|^2} \left(e^{i\Delta m \Delta t} \langle B_L | B_H \rangle + e^{-i\Delta m \Delta t} \langle B_H | B_L \rangle \right), \\
&= \frac{1}{4|p|^2} \left(e^{-\Gamma_L \Delta t} + e^{-\Gamma_H \Delta t} \right) + \frac{|p|^2 - |q|^2}{2|p|^2} e^{-\Gamma \Delta t} \cos(\Delta m \Delta t), \tag{5.32}
\end{aligned}$$

where Γ and Δm are defined in (5.29) and (5.30). Similarly, when the initial state is $|\bar{B}^0\rangle$, we find

$$\langle \bar{B}^0(\Delta t) | \bar{B}^0(\Delta t) \rangle = \frac{1}{4|q|^2} \left(e^{-\Gamma_L \Delta t} + e^{-\Gamma_H \Delta t} \right) - \frac{|p|^2 - |q|^2}{2|q|^2} e^{-\Gamma \Delta t} \cos(\Delta m \Delta t). \tag{5.33}$$

To clarify the meaning of (5.32), we insert the identity

$$\hat{1} = |B^0\rangle \langle B^0| + |\bar{B}^0\rangle \langle \bar{B}^0| \tag{5.34}$$

into the left hand side of (5.32) to obtain

$$\langle B^0(\Delta t) | B^0(\Delta t) \rangle = |\langle B^0 | B^0(\Delta t) \rangle|^2 + |\langle \bar{B}^0 | B^0(\Delta t) \rangle|^2. \tag{5.35}$$

This shows that the equation (5.32) provides the probability of remaining either as B^0 or \bar{B}^0 without collapsing until the time Δt after the particle is tagged as B^0 at the initial time. Note that the result (5.33) is also obtained from (5.32) by the interchange $p \leftrightarrow q$ as we noted before.

Taking the normalization condition of probability into account, we observe that the time distribution $P(\Delta t)$ expressing the probability density of decay at time Δt reads

$$\begin{aligned}
P(\Delta t) &= -\frac{d}{d\Delta t'} \langle B^0(\Delta t') | B^0(\Delta t') \rangle |_{\Delta t'=\Delta t} \\
&= \frac{1}{4|p|^2} \left(\Gamma_L e^{-\Gamma_L \Delta t} + \Gamma_H e^{-\Gamma_H \Delta t} \right) + \frac{|p|^2 - |q|^2}{2|p|^2} e^{-\Gamma \Delta t} \left(\Delta m \sin(\Delta m \Delta t) + \Gamma \cos(\Delta m \Delta t) \right). \tag{5.36}
\end{aligned}$$

The effective lifetime of this particle can then be evaluated as

$$\begin{aligned}
\tau_{\text{eff}}(B^0) &= \int_0^\infty d\Delta t' \Delta t' P(\Delta t') \\
&= - \int_0^\infty d\Delta t' \Delta t' \left[\frac{d}{d\tau} \langle B^0(\tau) | B^0(\tau) \rangle \right]_{\tau=\Delta t'} \\
&= - \left[\Delta t' \langle B^0(\Delta t') | B^0(\Delta t') \rangle \right]_{\Delta t'=0}^{\Delta t'=\infty} + \int_0^\infty d\Delta t' \langle B^0(\Delta t') | B^0(\Delta t') \rangle \\
&= \frac{1}{4|p|^2} \left(\frac{1}{\Gamma_L} + \frac{1}{\Gamma_H} \right) + \frac{|p|^2 - |q|^2}{2|p|^2} \frac{\Gamma}{\Gamma^2 + (\Delta m)^2}. \tag{5.37}
\end{aligned}$$

For the case of B mesons, the second term in (5.37) can be neglected when compared to the first term (the former is roughly one millionth of the latter).

5.2.3 Time distribution with postselection

Now we wish to implement the postselection and see how it affects the time distribution obtained before without the postselection. The postselected state can in general be written by a superposition of the B mesons as

$$|B_{\text{decay}}\rangle = r|B^0\rangle + s|\bar{B}^0\rangle, \quad (5.38)$$

where r, s are complex numbers satisfying the normalization condition,

$$|r|^2 + |s|^2 = 1. \quad (5.39)$$

When investigating the CP violation, the CP conjugate of the state in (5.38),

$$|\bar{B}_{\text{decay}}\rangle = s|B^0\rangle + r|\bar{B}^0\rangle, \quad (5.40)$$

will be used.

Before evaluating the time distribution with postselection, we first obtain the probability of finding the postselected state $|B_{\text{decay}}\rangle$ specified in (5.38) when the state under consideration is the B meson state $|B^0(\Delta t)\rangle$ given in (5.26) as

$$\begin{aligned} |\langle B_{\text{decay}}|B^0(\Delta t)\rangle|^2 &= |r^*\langle B^0|B^0(\Delta t)\rangle + s^*\langle \bar{B}^0|B^0(\Delta t)\rangle|^2 \\ &= \left| \left(\frac{r^*}{2} + \frac{qs^*}{2p} \right) e^{-\frac{\Gamma_L}{2}\Delta t} e^{-i\Delta tm_L} + \left(\frac{r^*}{2} - \frac{qs^*}{2p} \right) e^{-\frac{\Gamma_H}{2}\Delta t} e^{-i\Delta tm_H} \right|^2. \end{aligned} \quad (5.41)$$

Introducing the relative phase parameters φ and θ through

$$\frac{p}{q} = \frac{|p|}{|q|} e^{i\varphi}, \quad \frac{r}{s} = \frac{|r|}{|s|} e^{i\theta}, \quad (5.42)$$

we find

$$\begin{aligned} |\langle B_{\text{decay}}|B^0(\Delta t)\rangle|^2 &= \left(\frac{|r|^2}{2} + \frac{|q|^2|s|^2}{2|p|^2} \right) \frac{e^{-\Gamma_L\Delta t} + e^{-\Gamma_H\Delta t}}{2} \\ &\quad + \left(\frac{|r|^2}{2} - \frac{|q|^2|s|^2}{2|p|^2} \right) e^{-\Gamma\Delta t} \cos(\Delta m\Delta t) \\ &\quad + \frac{|r||q||s|}{|p|} \cos(\theta - \varphi) \frac{e^{-\Gamma_L\Delta t} - e^{-\Gamma_H\Delta t}}{2} \\ &\quad - \frac{|r||q||s|}{|p|} \sin(\theta - \varphi) e^{-\Gamma\Delta t} \sin(\Delta m\Delta t). \end{aligned} \quad (5.43)$$

The corresponding CP conjugate $|\langle \bar{B}_{\text{decay}}|\bar{B}^0(\Delta t)\rangle|^2$ is then obtained simply by putting $|p| \leftrightarrow |q|$ and $\varphi \leftrightarrow -\varphi$ in (5.43).

It is notable that either the second or the fourth term in (5.43) is non-vanishing for any combinations of the two parameters (r, s) specifying the postselection except the case $|p| = |q|$. These terms have oscillatory property which can be controlled by the parameters $|r|$ and θ . In our measurements, a certain set of particles generated by the decay of B mesons is measured properly in order to achieve the postselection specified by the state $|B_{\text{decay}}\rangle$ in (5.38).

The time distribution with postselection is then defined by

$$P(\Delta t|B^0 \rightarrow B_{\text{decay}}) = \frac{|\langle B_{\text{decay}}|B^0(\Delta t)\rangle|^2}{\int_0^\infty d\Delta t' |\langle B_{\text{decay}}|B^0(\Delta t')\rangle|^2}, \quad (5.44)$$

which is the conditional time distribution with B^0 at the initial time and B_{decay} at the time of decay. Similarly, its CP conjugate ($\bar{B}^0 \rightarrow \bar{B}_{\text{decay}}$) version is also defined by

$$P(\Delta t | \bar{B}^0 \rightarrow \bar{B}_{\text{decay}}) = \frac{|\langle \bar{B}_{\text{decay}} | \bar{B}^0(\Delta t) \rangle|^2}{\int_0^\infty d\Delta t' |\langle \bar{B}_{\text{decay}} | \bar{B}^0(\Delta t') \rangle|^2}. \quad (5.45)$$

Using (5.43), we obtain the denominator of (5.44) as

$$\begin{aligned} \int_0^\infty d\Delta t' |\langle B_{\text{decay}} | B^0(\Delta t') \rangle|^2 &= \left(\frac{|r|^2}{2} + \frac{|q|^2|s|^2}{2|p|^2} \right) \left(\frac{1}{2\Gamma_L} + \frac{1}{2\Gamma_H} \right) \\ &+ \left(\frac{|r|^2}{2} - \frac{|q|^2|s|^2}{2|p|^2} \right) \frac{\Gamma}{\Gamma^2 + (\Delta m)^2} \\ &+ \frac{|r||q||s|}{|p|} \cos(\theta - \varphi) \left(\frac{1}{2\Gamma_L} - \frac{1}{2\Gamma_H} \right) \\ &- \frac{|r||q||s|}{|p|} \sin(\theta - \varphi) \frac{\Delta m}{\Gamma^2 + (\Delta m)^2}. \end{aligned} \quad (5.46)$$

In particular, when the decay widths of the two mass eigenstates of the meson are considered to be equivalent, $|\Gamma_L - \Gamma_H|/\Gamma \ll 1$ (according to [57], $|\Gamma_L - \Gamma_H|/\Gamma < 0.01$), the time distribution (5.44) admits the form,

$$\begin{aligned} &P(\Delta t | B^0 \rightarrow B_{\text{decay}}) \\ &= \Gamma e^{-\Gamma \Delta t} \frac{\left(\frac{1}{2} + \frac{|q|^2(1-|r|^2)}{2|p|^2|r|^2} \right) + \left(\frac{1}{2} - \frac{|q|^2(1-|r|^2)}{2|p|^2|r|^2} \right) \cos(\Delta m \Delta t) - \frac{|q|\sqrt{1-|r|^2}}{|p||r|} \sin(\theta - \varphi) \sin(\Delta m \Delta t)}{\left(\frac{1}{2} + \frac{|q|^2(1-|r|^2)}{2|p|^2|r|^2} \right) + \left(\frac{1}{2} - \frac{|q|^2(1-|r|^2)}{2|p|^2|r|^2} \right) \frac{\Gamma^2}{\Gamma^2 + (\Delta m)^2} - \frac{|q|\sqrt{1-|r|^2}}{|p||r|} \sin(\theta - \varphi) \frac{\Gamma \Delta m}{\Gamma^2 + (\Delta m)^2}}. \end{aligned} \quad (5.47)$$

5.3 Weak value amplification

It is clear from the discussion so far that the influence of post-selection appears on the time distribution. However, it is difficult to understand from the viewpoint that amplification occurs due to the characteristics of weak values. Therefore, the average output, that is, the average survival time of particles, is considered in comparison with the case of normal weak value amplification.

Having seen that different time distributions can be obtained depending on whether the postselection was performed or not, we now evaluate the effective lifetime of the B meson with the postselection.

Actually, the mean value of the time distribution, which is the effective lifetime with the condition $B^0 \rightarrow B_{\text{decay}}$, is obtained by

$$\begin{aligned} \tau_{\text{eff}}(B^0 \rightarrow B_{\text{decay}}) &= \int_0^\infty d\Delta t' \Delta t' P(\Delta t' | B^0 \rightarrow B_{\text{decay}}) \\ &= \frac{\int_0^\infty d\Delta t' \Delta t' |\langle B_{\text{decay}} | B^0(\Delta t') \rangle|^2}{\int_0^\infty d\Delta t' |\langle B_{\text{decay}} | B^0(\Delta t') \rangle|^2}. \end{aligned} \quad (5.48)$$

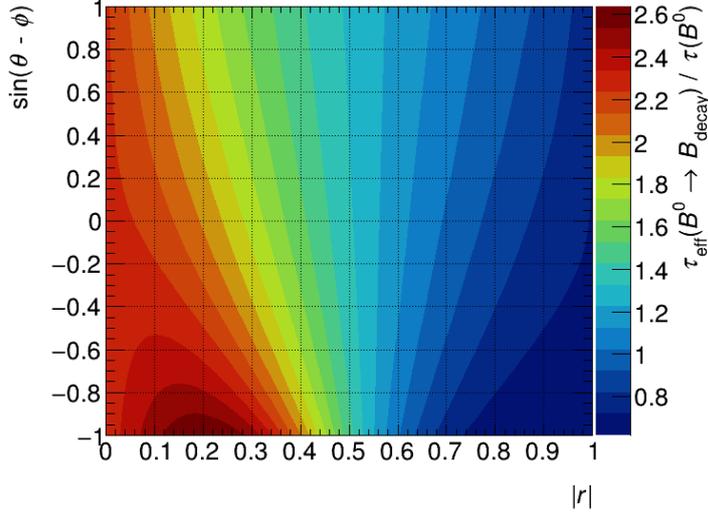


Figure 5.1: The ratio between the lifetime of B meson $\tau(B^0)$ and the effective lifetime (5.50). The effective lifetime can be extended about 2.6 times larger than $\tau(B^0)$ in this case when $|r| \sim 0.2$ and $\sin(\theta - \varphi) \sim -1$.

The denominator of (5.48) is the same as (5.46). The numerator is

$$\begin{aligned}
& \int_0^\infty d\Delta t' \Delta t' |\langle B_{\text{decay}} | B^0(\Delta t') \rangle|^2 \\
&= \left(\frac{|r|^2}{2} + \frac{|q|^2|s|^2}{2|p|^2} \right) \left(\frac{1}{2\Gamma_L^2} + \frac{1}{2\Gamma_H^2} \right) + \left(\frac{|r|^2}{2} - \frac{|q|^2|s|^2}{2|p|^2} \right) \frac{\Gamma^2 - (\Delta m)^2}{\{\Gamma^2 + (\Delta m)^2\}^2} \\
&+ \frac{|r||q||s|}{|p|} \cos(\theta - \varphi) \left(\frac{1}{2\Gamma_L^2} - \frac{1}{2\Gamma_H^2} \right) - \frac{|r||q||s|}{|p|} \sin(\theta - \varphi) \frac{2(\Delta m)\Gamma}{\{\Gamma^2 + (\Delta m)^2\}^2}. \quad (5.49)
\end{aligned}$$

When we have $|\Gamma_L - \Gamma_H|/\Gamma \ll 1$ and hence are allowed to put $\Gamma_L = \Gamma_H = \Gamma$ approximately, then the effective lifetime (5.48) admits the closed form,

$$\begin{aligned}
& \tau_{\text{eff}}(B^0 \rightarrow B_{\text{decay}}) \\
&= \frac{\left(1 + \frac{|q|^2|s|^2}{|p|^2|r|^2}\right) \frac{1}{\Gamma^2} + \left(1 - \frac{|q|^2|s|^2}{|p|^2|r|^2}\right) \frac{\Gamma^2 - (\Delta m)^2}{\{\Gamma^2 + (\Delta m)^2\}^2} + \frac{|q||s|}{|p||r|} \sin(\theta - \varphi) \frac{4(\Delta m)\Gamma}{\{\Gamma^2 + (\Delta m)^2\}^2}}{\left(1 + \frac{|q|^2|s|^2}{|p|^2|r|^2}\right) \frac{1}{\Gamma} + \left(1 - \frac{|q|^2|s|^2}{|p|^2|r|^2}\right) \frac{\Gamma}{\Gamma^2 + (\Delta m)^2} + \frac{|q||s|}{|p||r|} \sin(\theta - \varphi) \frac{2\Delta m}{\Gamma^2 + (\Delta m)^2}} \\
&= \frac{(1 + |A_w|^2) \frac{1}{\Gamma^2} + (1 - |A_w|^2) \frac{\Gamma^2 - (\Delta m)^2}{\{\Gamma^2 + (\Delta m)^2\}^2} + 4\text{Im}[A_w] \frac{(\Delta m)\Gamma}{\{\Gamma^2 + (\Delta m)^2\}^2}}{(1 + |A_w|^2) \frac{1}{\Gamma} + (1 - |A_w|^2) \frac{\Gamma}{\Gamma^2 + (\Delta m)^2} + 2\text{Im}[A_w] \frac{\Delta m}{\Gamma^2 + (\Delta m)^2}}, \quad (5.50)
\end{aligned}$$

where A_w is the weak value in (5.5), which now takes the form,

$$A_w = \frac{|q||s|}{|p||r|} \cos(\theta - \varphi) + i \frac{|q||s|}{|p||r|} \sin(\theta - \varphi). \quad (5.51)$$

Note that we have obtained the result (5.50) without resorting to the first-order approximation for $\Delta m/\Gamma$, which is actually required for us since $\Delta m/\Gamma \simeq 0.77$ for the B meson system we are considering [57, p.63]. For B mesons, Figure 5.1 shows that this effective lifetime can be 2.6 times the lifetime $\tau(B^0)$.

Returning to the general case, we note that when $\Delta m/\Gamma \ll 1$ holds, one may expand the effective lifetime (5.50) in terms of $\Delta m/\Gamma$ to obtain

$$\tau_{\text{eff}}(B^0 \rightarrow B_{\text{decay}}) = \frac{1}{\Gamma} + \frac{1}{\Gamma} \text{Im}[A_w] \frac{\Delta m}{\Gamma} + O((\Delta m/\Gamma)^2). \quad (5.52)$$

This is equal to (5.9).

This amplification effect can also be seen in the detection of CP violation. For simplicity, we assume $|p| = |q|$ and $\Delta m/\Gamma \ll 1$ so that the difference between the denominators of (5.47) is small enough. Then, the difference between their time distributions is

$$\begin{aligned} P(\Delta t|B^0 \rightarrow B_{\text{decay}}) - P(\Delta t|\bar{B}^0 \rightarrow \bar{B}_{\text{decay}}) \\ \propto -\frac{\sqrt{1-|r|^2}}{|r|} (\sin(\theta - \varphi) - \sin(\theta + \varphi)) e^{-\Gamma\Delta t} \sin(\Delta m\Delta t) \\ = 2\frac{\sqrt{1-|r|^2}}{|r|} \cos\theta \sin\varphi e^{-\Gamma\Delta t} \sin(\Delta m\Delta t). \end{aligned} \quad (5.53)$$

This result indicates that the parameters $|r|$ and θ introduced in our postselection may amplify the difference in the conditional time distribution.

5.4 Process of postselection and related issues

In order to obtain experimental results that correspond to the conditional time distributions derived above, it is necessary to clarify the measurement operations required further. In the actual experiment we estimate the postselected state of the B meson $|B_{\text{decay}}\rangle$ through measurements of various decayed particles. As a possible scenario for implementing the measurement we wish to achieve, we shall consider a decay mode that involves photons for which the postselection has been commonly discussed in optical systems (See, for example, [19]).

To this end, let us first expand the state (5.26) of the B meson immediately before the decay as

$$|B^0(\Delta t)\rangle = a(\Delta t)|B^0\rangle + b(\Delta t)|\bar{B}^0\rangle. \quad (5.54)$$

At this point, we recall that the Standard Model predicts that photons produced in the decay $b \rightarrow s\gamma$ are mainly left-handed [76]. For simplicity, in this paper, we assume that the involved photons are always left-handed, which implies that the photons produced from the decay of anti- b quarks are always right-handed. It then follows that the combinations of the decayed particles are

$$|B^0\rangle \leftrightarrow |K^{*0}\rangle|\gamma_R\rangle, \quad |\bar{B}^0\rangle \leftrightarrow |\bar{K}^{*0}\rangle|\gamma_L\rangle, \quad (5.55)$$

where $|\gamma_L\rangle$ and $|\gamma_R\rangle$ represent the state of the left-handed photon and that of the right-handed photon, respectively. This implies that, if we let \hat{U} be the unitary operator describing the time development connecting the state of the B meson and that of the decayed particles, and if we can assume that this process preserves the CP symmetry $[\mathcal{CP}, \hat{U}] = 0$, then we have

$$\begin{aligned} \hat{U}|B^0\rangle &= c|K^{*0}\rangle|\gamma_R\rangle + \dots, \\ \hat{U}|\bar{B}^0\rangle &= c|\bar{K}^{*0}\rangle|\gamma_L\rangle + \dots, \end{aligned} \quad (5.56)$$

with a common constant c , where \dots stands for other particles generated from the decay. This means that the state just after the decay can be written as

$$\hat{U}|B^0(\Delta t)\rangle = ca|K^{*0}\rangle|\gamma_R\rangle + cb|\bar{K}^{*0}\rangle|\gamma_L\rangle + \dots \quad (5.57)$$

Now, in order to implement the postselection $|B_{\text{decay}}\rangle$ in (5.38) for the state of the B meson by choosing a particular mode in the decay, the postselection for the decay mode must fulfill a certain consistency condition. To see this, let $|K_F^*\rangle$ and $|\gamma_F\rangle$ be the states we choose as postselected states of the K^* and γ systems, respectively. The condition that must be respected is that, as a function of the parameters r and s specifying the postselection (5.38), the transition probabilities evaluated for the postselected state of the decay mode be equivalent to that of the original B meson up to a constant, that is,

$$|\langle\gamma_F|\langle K_F^*|\hat{U}|B^0(\Delta t)\rangle|^2 \propto |\langle B_{\text{decay}}|B^0(\Delta t)\rangle|^2. \quad (5.58)$$

Expanding $|K_F^*\rangle$ in terms of the basis states as

$$|K_F^*\rangle = \xi_1|K^{*0}\rangle + \xi_2|\bar{K}^{*0}\rangle, \quad (5.59)$$

and similarly $|\gamma_F\rangle$ as

$$|\gamma_F\rangle = \eta_1|\gamma_R\rangle + \eta_2|\gamma_L\rangle, \quad (5.60)$$

and substituting these in (5.58), we obtain

$$\begin{aligned} & \langle K_F^*|\langle\gamma_F|\hat{U} \\ &= \left(\xi_1^*\eta_1^*\langle K^{*0}|\langle\gamma_R| + \xi_2^*\eta_1^*\langle\bar{K}^{*0}|\langle\gamma_R| + \xi_1^*\eta_2^*\langle K^{*0}|\langle\gamma_L| + \xi_2^*\eta_2^*\langle\bar{K}^{*0}|\langle\gamma_L| \right) \hat{U} \\ & \quad \times (|B^0\rangle\langle B^0| + |\bar{B}^0\rangle\langle\bar{B}^0|) \\ &= c(\xi_2^*\eta_2^*\langle\bar{B}^0| + \xi_1^*\eta_1^*\langle B^0|), \end{aligned} \quad (5.61)$$

where $|B^0\rangle\langle B^0| + |\bar{B}^0\rangle\langle\bar{B}^0| = \hat{1}$ has been used. This implies that the consistency condition (5.58) is satisfied if we choose the expanding parameters as

$$\frac{\xi_1\eta_1}{\xi_2\eta_2} = \frac{r}{s}. \quad (5.62)$$

We thus learn that there are a variety of possible postselections on the state of the decay mode fulfilling our condition (5.58).

One particular postselection is provided by choosing the CP eigenstate $|K_1^{*0}\rangle$ for $|K_F^*\rangle$ in (5.59), namely,

$$|K_F^*\rangle = |K_1^{*0}\rangle = \frac{1}{\sqrt{2}} (|K^{*0}\rangle + |\bar{K}^{*0}\rangle). \quad (5.63)$$

This CP eigenstate $|K_1^{*0}\rangle$ is experimentally distinguished as it admits the same decay mode $K^* \rightarrow K_S^0\pi^0$ and hence its choice, including the phase, can be considered to be relatively easy compared to other possibilities. With this choice, one easily finds from (5.62) the other parameters to obtain the postselection of the photon state (5.60) as

$$|\gamma_F\rangle = r|\gamma_R\rangle + s|\gamma_L\rangle. \quad (5.64)$$

This postselection for photons may be realized by the standard method used for photonic systems provided that the range of energy can be extended considerably. Summing up, we

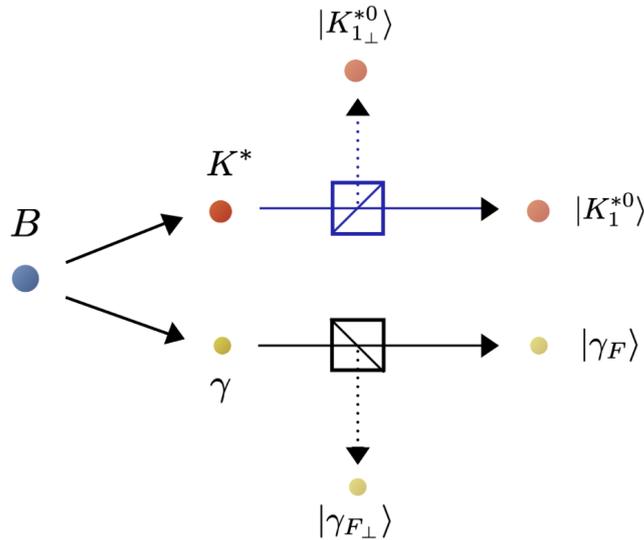


Figure 5.2: After the decay of the B meson, the selection is performed for each of K^* and γ . For the K^* , it is measured whether it is $|K_1^{*0}\rangle$ given by (5.63) or $|K_{1\perp}^{*0}\rangle$, which is the orthogonal state. Similarly, for the γ , we measure whether it is $|\gamma_F\rangle$ or its orthogonal state $|\gamma_{F\perp}\rangle$. We evaluate the time distribution only the decay events in which each measurement result agrees with its selected state. In the case of low-energy photons (visible light) often used in weak measurements, the beam splitter is commonly used for the postselection along with the wave plate and the compensation plate.

have learned that, although the actual postselection is performed on the particles which are generated at the time of the decay, we can consider it as the postselection on the particle in effect, and there are certainly many possible choices to achieve this goal.

Incidentally, we note that, even if we start with the CP conjugate state,

$$|\bar{B}^0(\Delta t)\rangle = a(\Delta t)|\bar{B}^0\rangle + b(\Delta t)|B^0\rangle, \quad (5.65)$$

instead of (5.54), we still can repeat the above argument with the CP operation, reaching essentially the same result (5.58) with $|B^0(\Delta t)\rangle$ replaced by $|\bar{B}^0(\Delta t)\rangle$.

Note also that K^* and γ produced by the decay of the B meson evolve over time after the production. While the state of the photon does not change as long as it flies in the vacuum, the K^* state changes during the flight. Therefore, an additional operation is required to specify the state (5.63) taking into account this time evolution.

5.5 Issues related to experimental method and systems

In a later section, we introduce a numerical experiment that examines the experimental error due to weak value amplification, but in this simulation, the calculation is performed with the Belle II experiment operated at High Energy Accelerator Research Organization (KEK) in mind. In this section, we introduce discussions related to referring to this experiment.

We shall now move on to discussing the weak measurement of the $B^0 - \bar{B}^0$ mixing using the $B^0 \rightarrow K^{*0}\gamma$ decay channel from the experimental point of view.

5.5.1 Extension of time distribution in the negative direction

In the Belle II experiment, B mesons are generated through the decay of $\Upsilon(4S)$ particles. $\Upsilon(4S)$ particle is a kind of particles produced by the collision of electrons and positrons. In our discussion, we shall restrict ourselves to the type of flavor tagging carried out by identifying one of the species in the ($\mathcal{C} = -1$) entangled pair of B mesons as adopted in the Belle II experiment.

This requires us to take into account negative Δt as well as positive ones, since the decay of one meson in the entangled pair may have already occurred when the tagging of the other meson is made. To accommodate this extension in the probability density function $P(\Delta t|B^0 \rightarrow B_{\text{decay}})$ in (5.47), we consider, in addition to the transition probability $|\langle B_{\text{decay}}|B^0(\Delta t)\rangle|^2$ in (5.41) for $\Delta t > 0$, the corresponding one

$$|\langle B^0|B_{\text{decay}}(-\Delta t)\rangle|^2 = |\langle B^0|e^{i\Delta t\hat{H}}|B_{\text{decay}}\rangle|^2 \quad \text{for } \Delta t < 0. \quad (5.66)$$

This implies $\hat{H} \rightarrow \hat{H}^\dagger$ when we go over from $\Delta t > 0$ to $\Delta t < 0$, and for our case where the Hamiltonian admits the form (5.28). This amounts to the parameter replacement $(p, q, \Gamma, \Delta\Gamma) \rightarrow (q^*, p^*, -\Gamma, -\Delta\Gamma)$.

In the probability density function $P(\Delta t|B^0 \rightarrow B_{\text{decay}})$ in (5.47), the effect of the CP violation arises from the difference between p and q . In the SM, the difference of their absolute scalars $1 - |q|/|p|$ is expected to be small $\mathcal{O}(10^{-3})$ [57, p.63] in the $B^0 - \bar{B}^0$ system. Then, assuming $|q|/|p| = 1$ for simplicity, we find from (5.47) and its CP conjugate partner that the probability density functions reduce to

$$P_{\text{phys}}(\Delta t|B^0 \rightarrow B_{\text{decay}}) = \frac{e^{-\frac{|\Delta t|}{\tau}}}{2N} \left(1 + (2|r|^2 - 1) \cos(\Delta m\Delta t) - 2|r|\sqrt{1 - |r|^2} \sin(\theta - \varphi) \sin(\Delta m\Delta t) \right), \quad (5.67)$$

$$P_{\text{phys}}(\Delta t|\bar{B}^0 \rightarrow \bar{B}_{\text{decay}}) = \frac{e^{-\frac{|\Delta t|}{\tau}}}{2N} \left(1 + (2|r|^2 - 1) \cos(\Delta m\Delta t) - 2|r|\sqrt{1 - |r|^2} \sin(\theta + \varphi) \sin(\Delta m\Delta t) \right), \quad (5.68)$$

where $\tau = 1/\Gamma$ is the B^0 lifetime (which is 1.519 ps [57, p.63]) and

$$N = \tau \left(1 + \frac{2|r|^2 - 1}{1 + (\Delta m\tau)^2} \right), \quad (5.69)$$

is a normalization factor. (The formula (5.68) is obtained from (5.67) by putting $\varphi \rightarrow -\varphi$, which is a consequence of the interchange $p \leftrightarrow q$. The extension to $\Delta t < 0$ in this simplified case is achieved just by putting $\tau \rightarrow -\tau$ according to the note in the previous paragraph, resulting in the exponential factor in (5.67) and (5.68).) We note that in these formulae the only measurable parameter is φ which corresponds to $2\phi_1$ in the SM (see Section 12 in [57]), where ϕ_1 is one of the interior angles of the unitarity triangle. If a new phase arises from new physics in $B \rightarrow K^*\gamma$ penguin diagram, the measured φ is deviated from the $2\phi_1$.

5.5.2 Time measurement method

All the theoretical discussions so far are on the time evolution of particles in a stationary system. In high-energy experiments, the particles move relativistically, and the time from the flavor tag to the decay of the particle is obtained from the distance traveled by the particle.

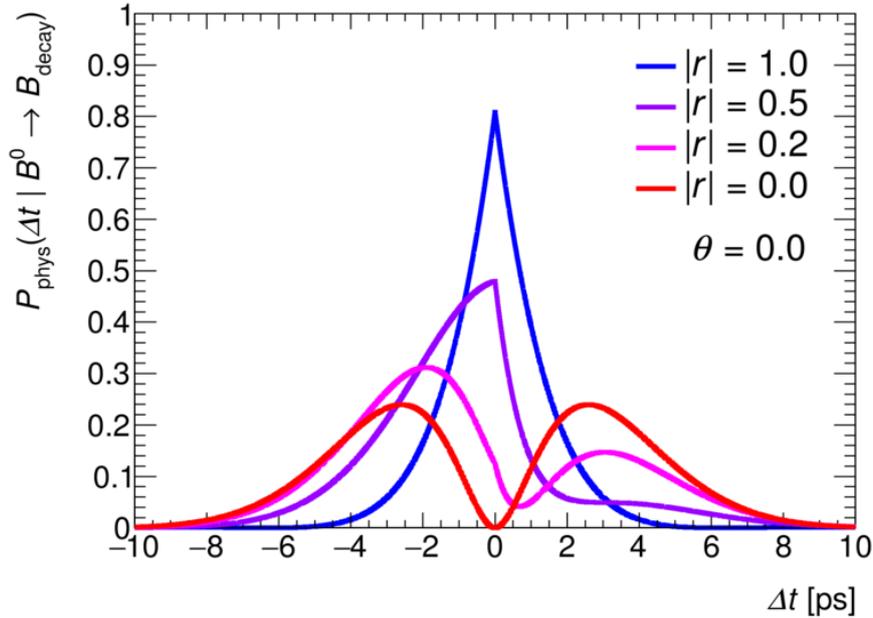


Figure 5.3: The time distribution theoretically expected from (5.67). The distribution actually obtained in the experiment is subject to changes due to noise mixing and specific measurement operations.

The time when the observer performs the flavor tag is selected as the reference for the coordinates (z, t) where the observer is and the coordinates (z', t') fixed to the particle.

The decay time distribution calculated in the above discussion so far refers to the time at the coordinates fixed to the particle. In the case that the particle decay has occurred at $(z', t') = (0, \Delta t)$, the corresponding z is

$$z = \gamma(z' + vt') = \gamma v \Delta t. \quad (5.70)$$

The B meson decaying to the f_{rec} final state, $K^{*0}\gamma$ in this study, is reconstructed from the decay particles detected by the Belle II detector. The remaining particles in the detector assumed as the particles from the other B meson decaying to the f_{tag} final state. They are used to tag the flavor of the B meson decayed to the f_{rec} state, where the tagging time t_{tag} is the instant when the decay of the B meson into f_{tag} occurred. The probability density functions of B and \bar{B} mesons, given by (5.47) and its CP conjugate partner, describe the time variation of the decays as a function of Δt which can be determined from the displacement between f_{rec} and f_{tag} vertices along with z -axis of the beam, thanks to the asymmetric electron-positron beam energy.

The SuperKEKB is an asymmetric energy electron-positron collider with 7 GeV and 4 GeV of electron and positron beam energy, respectively. These electron and positron collide. Then, a particle $\Upsilon(4S)$ is generated. This particle is Lorentz-boosted with $\beta\gamma = 0.28$ along the direction of the electron beam.

The Belle II detector consists of a number of sub-detectors surrounding the electron-positron collision point. The generated B meson repeats further decay. By finding the intersection of the tracks traced by the particles generated by such decay, the position where the decay has occurred can be estimated. In addition, the decay time can be estimated through the estimation of such a position. Of course, the decay time obtained by this estimation includes an error. Subsection 5.6.1 describes how to deal with such errors.

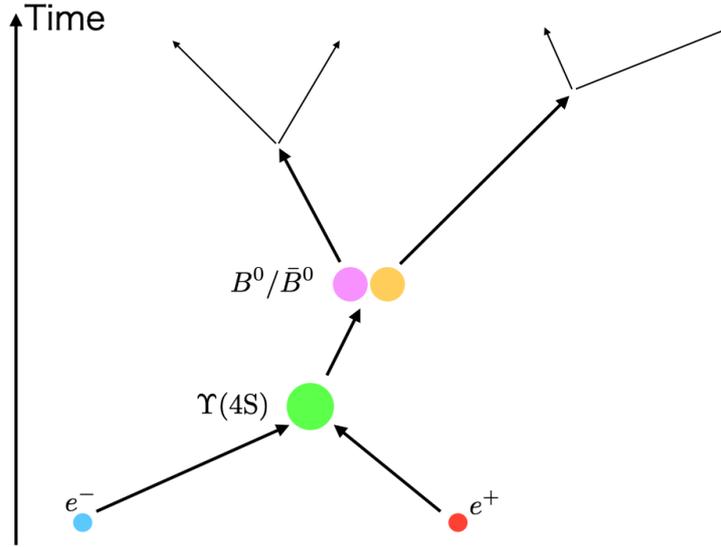


Figure 5.4: In the Belle II experiment, the SuperKEKB accelerator accelerates electrons to 7 GeV for energy and positrons to 4 GeV, causing them to collide. Then, it decays into a pair of B meson and anti- B meson via $\Upsilon(4S)$ particle. The survival time is estimated from the flight distance of the particles.

5.6 Uncertainty evaluation with Monte-Carlo simulation

Even in high-energy experiments, it is necessary to comprehensively examine the measurement conditions and measurement environment to see if performing the WVA has the effect of reducing the experimental uncertainty. In order to show this, we have considered performing a simulation to evaluate the expected uncertainty if the experiment of the present proposal can be realized, referring to the measurement conditions of the accelerator experiment, particularly, the Belle II experiment operated at KEK.

It should be noted that this simulation is carried out without considering the details of the actual preselection and postselection method because the corresponding operation is unknown in high energy experiments. Therefore, the true error in the actual experiment needs to be examined again when the actual scheme of the postselection is determined, and the result in this section should be regarded as showing only that the error can be changed by selecting the postselected state.

This Monte-Carlo simulation was carried out by Dr. Higashino in cooperation with Dr. Takubo, Prof. Higuchi, and Prof. Ishikawa.

5.6.1 Event simulation

To perform a Monte-Carlo simulation, we have to obtain a ‘true’ probability distribution and determine the number of trials. The probability distribution is expected to match the distribution obtained by repeating the actual experiment a sufficiently large number of times. Such a true distribution is not determined solely by the physical phenomenon targeted by the experiment, but it is necessary to comprehensively examine the signal corresponding to noise and the performance and characteristics of the experimental device.

The effect of noise also depends on the scheme of the postselection, so it is difficult to fully model this at this time. Also, the decrease in the number of statistics due to the postselection is estimated to be about a factor in this case, and in a certain sense that

the mode is selected in the experiments that have already been performed. Therefore, we assume that the number of statistics is constant and the signal-to-noise ratio is also constant. We should be careful of this point when we compare our results with other results of simulations or experiments.

Trial number

Improvement of sensitivity to the CP violation with weak measurement is evaluated in $B^0 \rightarrow K^{*0}\gamma$ decay as a benchmark by using pseudo-experiments with toy Monte-Carlo samples.

The experiment will collect $550 \times 10^8 B\bar{B}$ pairs at 50 ab^{-1} , that are produced from resonance decays of $\Upsilon(4S)$ at the SuperKEKB. The number of events is assumed to be $550 \times 10^8 B\bar{B}$ in accordance with the Belle II experiment, where $B\bar{B}$ contains both $B^0\bar{B}^0$ and B^+B^- . The branching fractions of $\Upsilon(4S) \rightarrow B^0\bar{B}^0$ and $B^0 \rightarrow K^{*0}\gamma$ are considered as 0.49 [57, p.87] and 4.2×10^{-5} [57, p.68], respectively.

The efficiency to identify the flavor of B mesons, B^0 or \bar{B}^0 , decaying into f_{tag} is assumed to be 0.136, where the wrong tagging fraction is 0.02 in the Belle experiment [77].

The reconstruction and selection efficiency will also be important to the results in the Belle experiment, which is estimated to be 2.1% [78]. Finally, the signal yield is expected to be 3.3×10^3 in the SM prediction.

Resolution of the measurement of the decay time

The time resolution is determined by various sources related to the position resolution of decay vertices. The causes and classifications of such errors are similar to those treated as intractable errors in the previous chapter, but it is assumed that the errors are also characterized by the same probability distribution in each measurement. That is, if any probability distribution $P(x)$ is given as the probability distribution for the true value, the distribution function $\tilde{P}(x)$, which is the convergence destination of distribution obtained by repeating actual measurements many times, can be written with an appropriate function $R(x)$ as follows.

$$\tilde{P}(x) = \int_{-\infty}^{\infty} d\tilde{x} P(\tilde{x}) R(x - \tilde{x}) \quad (5.71)$$

The time resolution functions for the signal R_{phys} and background R_{bkg} are evaluated separately and convoluted to each distribution, while the same profile is used both for R_{phys} and R_{bkg} . The time resolution depends on the position resolution of decay vertices which should be considered to reproduce the timing response in the actual experiment. Figure 5.5 shows the residual of Δt_{sim} and Δt_{true} , where Δt_{sim} is the simulated Δt where the position resolution of decay vertices is taken into account with the Belle experiment [78], and Δt_{true} is the true value of Δt without the detector response.

To obtain the resolution function of Δt , Figure 5.5 is fitted by an empirical function, called ‘double-sided crystal ball function’. The double-sided crystal ball function consists of a combination of a Gaussian and power law tails, defined as

$$R(x) = N_{\text{res}} \times \begin{cases} e^{-\frac{x^2}{2}}, & \text{for } -\alpha_L \leq X \leq \alpha_H, \\ e^{-\frac{\alpha_L^2}{2}} \left[\frac{\alpha_L}{n_L} \left(\frac{n_L}{\alpha_L} - \alpha_L - X \right) \right]^{-n_L}, & \text{for } X < -\alpha_L, \\ e^{-\frac{\alpha_H^2}{2}} \left[\frac{\alpha_H}{n_H} \left(\frac{n_H}{\alpha_H} - \alpha_H + X \right) \right]^{-n_H}, & \text{for } X > \alpha_H, \end{cases} \quad (5.72)$$

with the shorthand,

$$X = \frac{x - \mu}{\sigma}, \quad (5.73)$$

where N_{res} is the normalization factor, μ and σ are the mean and width of the Gaussian, α_H and α_L are the positions of transition from the Gaussian to the power low tails on the higher or lower sides, and n_H and n_L are the exponents of the high and low tails, respectively.

Event generator

Taking all these factors into account, we learn that the generation of simulated events follows the probability density function,

$$\begin{aligned} \mathcal{P}(\Delta t | B^0 \rightarrow B_{\text{decay}}) = \int_{-\infty}^{\infty} d(\Delta t') & \left[f_{\text{phys}} P_{\text{phys}}(\Delta t' | B^0 \rightarrow B_{\text{decay}}) R_{\text{phys}}(\Delta t - \Delta t') \right. \\ & \left. + (1 - f_{\text{phys}}) P_{\text{bkg}}(\Delta t') R_{\text{bkg}}(\Delta t - \Delta t') \right], \end{aligned} \quad (5.74)$$

and its CP conjugate partner $\mathcal{P}(\Delta t | \bar{B}^0 \rightarrow \bar{B}_{\text{decay}})$ in the corresponding form, both of which are characterized by the same fraction of the signal f_{phys} taken from [78], together with P_{phys} and P_{bkg} which are the distributions of the signal equivalent to (5.67) and the background, respectively. The main background sources are the light quark pair production process ($e^+e^- \rightarrow q\bar{q}$ with $q = u, d, s, c$) and the $e^+e^- \rightarrow B\bar{B}$ process where the B meson decays into final states different from the signal leading to misidentification. The profile of the background distribution is empirically determined from the results in the Belle experiment as $P_{\text{bkg}}(\Delta t') = \frac{1}{\tau_{\text{bkg}}} e^{-|\Delta t'|/\tau_{\text{bkg}}}$, where τ_{bkg} is set to be 0.896 ps.

Figure 5.5 shows the residual of Δt_{sim} and Δt_{true} , where Δt_{sim} is the simulated Δt where the position resolution of decay vertices is taken into account with the Belle experiment [78], and Δt_{true} is the true value of Δt without the detector response. The resolution functions are well modeled by fitting the simulated data under the parameters shown there.

The postselection parameters r and θ in (5.67) and (5.68) represent the mixing ratio and the phase difference between $|B^0\rangle$ and $|\bar{B}^0\rangle$ states. They are the parameters of the postselection that have to be determined by event selections in the experiment. Although their determination is an important technical subject in the actual experiment, for simplicity we assume that we have some means to determine them as we wish. The selection efficiencies of r and θ are assumed as 50% and 100%, respectively. Then, the number of signal events turns out to be 1.7×10^3 with the expected number of background events 0.9×10^3 .

5.6.2 Signal extraction and uncertainty in this analysis

We use the maximum likelihood method to estimate the CP parameter φ . We also evaluate the estimation accuracy of the CP parameter φ using Neyman's procedure shown in Section 2.4.1.

The probability density function given in (5.74) depends on the fitting parameter φ as well as systematic uncertainties coming from the three sources: the background

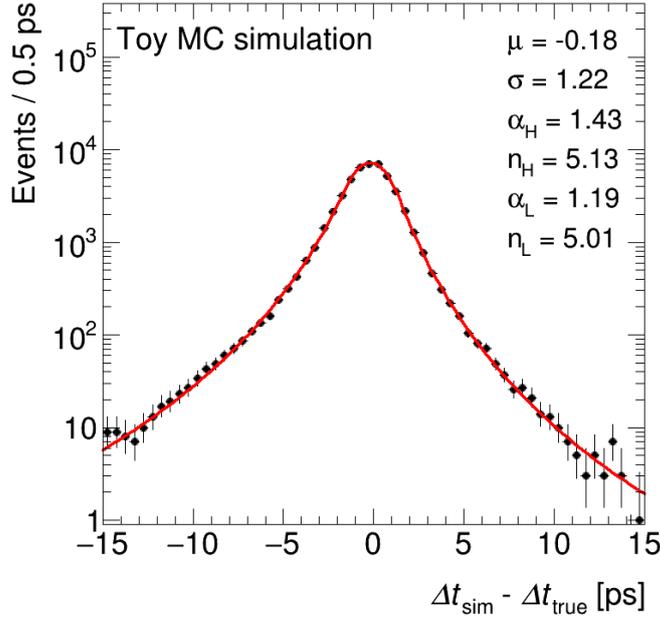


Figure 5.5: Distribution of the residual of Δt_{sim} and Δt_{true} . The dots are simulated data. The red line is the fitted function which is empirically modeled by the double-sided crystal ball function with parameters shown above.

estimation, the timing scale, and the timing resolution, which correspond to f_{phys} , μ and σ in (5.74) with (5.73).

A set of systematic uncertainties is parameterized by the so-called nuisance parameters θ^{nuis} . To consider the effect of the systematic uncertainties as well as statistical uncertainty, we use the likelihood function L

$$L = \mathcal{P}(\Delta t, \theta_{f_{\text{phys}}}^{\text{nuis}}, \theta_{\mu}^{\text{nuis}}, \theta_{\sigma}^{\text{nuis}}) G(\theta_{f_{\text{phys}}}^{\text{nuis}}) G(\theta_{\mu}^{\text{nuis}}) G(\theta_{\sigma}^{\text{nuis}}), \quad (5.75)$$

where these $G(\theta^{\text{nuis}})$ are all given by the Gaussian distribution

$$G(\theta^{\text{nuis}}) = \frac{1}{\sqrt{2\pi}\sigma_{\text{nuis}}} e^{-\frac{(\theta^{\text{nuis}} - \mu_{\text{nuis}})^2}{2\sigma_{\text{nuis}}^2}}, \quad (5.76)$$

which provides constraints on the systematic uncertainties [79]. We call the uncertainty calculated by the likelihood function (5.75) the total uncertainty. The parameter σ_{nuis} is the standard deviation and μ_{nuis} is the expectation value of the parameter θ^{nuis} . (see Section 2.4 and Section 40 in [57]). In our simulation, we fix these parameters σ_{nuis} referring to the value used in analysis of the Belle experiment [78, 80].

On the other hand, when the nuisance parameters are fixed to the expectation values μ_{nuis} , that is, when we use the probability distribution function

$$\tilde{L} = \mathcal{P}(\Delta t, \theta_{f_{\text{phys}}}^{\text{nuis}}, \theta_{\mu}^{\text{nuis}}, \theta_{\sigma}^{\text{nuis}}), \quad (5.77)$$

as the likelihood function, the uncertainty calculated by Neyman's procedure should be purely statistical.

The maximum likelihood fit is performed simultaneously for the distributions of $B^0 \rightarrow B_{\text{decay}}$ and $\bar{B}^0 \rightarrow \bar{B}_{\text{decay}}$. The dominant systematic uncertainties derive from the background estimation, the timing scale, and the resolution of uncertainties in the position of the reconstructed vertices. Only these uncertainties are considered in the maximum likelihood fit.

5.7 Analysis result

A pseudo-experiment is carried out and repeated a thousand times to estimate the measurement precision of φ par particular set of the postselection parameters $|r|$ and θ , in which events are randomly generated, following the probability density functions, (5.74) and its CP conjugate partner. These parameters are scanned in the range of $0.1 \leq |r| \leq 0.9$ with the intervals of 0.1 and similarly in the range of $-180 \leq \theta \leq 180$ degrees with the intervals of 36 degrees. The initial value of φ is set to 44.4 degrees which corresponds to $\varphi = 2\phi_1$ with the world average value of $\phi_1 = 22.2 \pm 0.7$ degrees [81]. The unbinned maximum likelihood fit is performed to each pseudo-experiment.

Figure 5.6 shows the Δt distributions when $|r|$ and θ are selected to be 0.5 and 0.0 degree, respectively. The fitted functions are shown in the plot, which are normalized to the number of events in each distribution. The shift of the peak position between $B^0 \rightarrow K^{*0}\gamma$ and $\bar{B}^0 \rightarrow K^{*0}\gamma$ shows the effect of the CP violation in the $B^0 - \bar{B}^0$ mixing with the CP parameter φ .

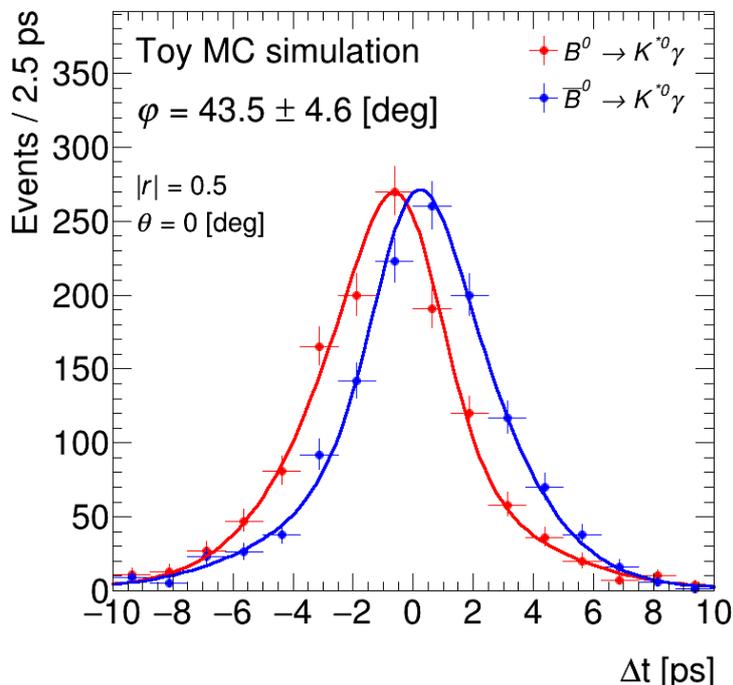


Figure 5.6: Distribution of Δt for B^0 and \bar{B}^0 initial states. The postselection parameters $|r|$ and θ are selected to be 0.5 and 0.0 degrees, respectively. The red and blue lines represent the best fit probability density functions, $\mathcal{P}(\Delta t | B^0 \rightarrow B_{\text{decay}})$ and $\mathcal{P}(\Delta t | \bar{B}^0 \rightarrow \bar{B}_{\text{decay}})$.

Figure 5.7 shows the total uncertainty of measured φ as a function of $|r|$ and θ which are used as the input values. This plot indicates that the measurement precision is improved by 20% with $|r| = 0.5$ compared to the case in which the final state is the CP eigenstate like $B^0 \rightarrow J/\psi K_S^0$ for which $|r| = \frac{1}{\sqrt{2}} \simeq 0.7$. Note that the precision is virtually insensitive to θ . The systematic uncertainty is in the range of 0.2 – 1.1 degrees while the statistical uncertainty varies from 4.7 to 13.2 degrees, depending on the postselection parameters, indicating the dominance of the statistical uncertainty in our analysis.

We should be careful of that it does not show that our method will immediately be a way to improve the accuracy of the current Belle II experiment. This 20% improvement

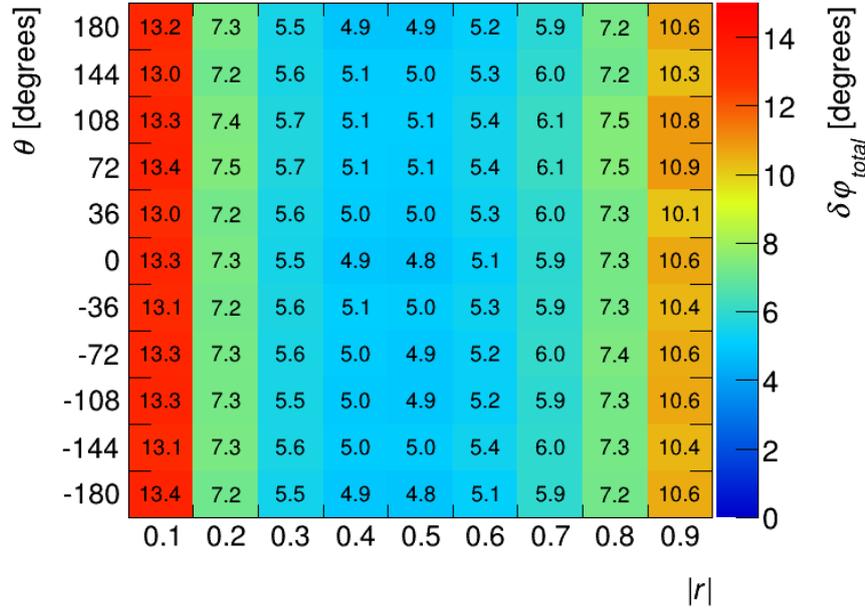


Figure 5.7: Total uncertainty of the φ measurement ($\delta\varphi_{\text{total}}$) as a function of $|r|$ and θ used as the input value.

of the accuracy requires comparing the results of the accuracy using the same number of statistics.

5.8 Comments on the time distribution and Monte-Carlo simulation

5.8.1 On the derivation of the time distribution

In the discussion of applying the weak value to the decay time distribution, it is notable that a prescription at calculating the conditional time distribution. The probability distribution we directly calculated is the probability that the state at the postselection is $|B_{\text{decay}}\rangle$ with a fixed time. Therefore, the time is a constant, and the postselected state $|B_{\text{decay}}\rangle$ is a ‘variable’ of the calculated probability distribution. However, the probability distribution actually obtained is a distribution in which the time is a variable of the distribution and the postselected state $|B_{\text{decay}}\rangle$ is fixed. That is, the constant and the variable of the probability distribution are swapped. This is also seen in Neyman’s procedure introduced in Chapter 2.

According to the Bayes’ theorem

$$P(A|B) = \frac{P(A)}{P(B)}P(B|A), \quad (5.78)$$

we could consider that this procedure can be justified simply by dealing with a value corresponding to

$$\frac{P(A)}{P(B)}. \quad (5.79)$$

In fact, we assume that the value (5.79) is constant. However, it is not obvious how to handle the value (5.79). In some cases, it may be necessary to prepare a different argument rather than considering it as a constant.

5.8.2 On the validity of the results on the estimated uncertainty

As already mentioned in Event generator in Subsection 5.6.1, we have assumed the same statistics to obtain the result (Figure 5.7), regardless of the choice of the variables r and θ that specify the postselected state. In the actual experiments, the CP -violating parameters are measured using modes with higher decay branching ratios so that the accuracy can be better than our result. To obtain the 20% accuracy improvement mentioned at the end of Section 5.7, we need to secure the number of statistics for all r and θ to the same extent for unbiased comparison.

If we want to choose r and θ in the experiment freely, we need to consider the efficiency of the postselection as well as the determination of the decay mode to be used in order to determine the statistical numbers. In our analysis, this efficiency is assumed to be fixed to 50%, but we cannot select the polarization of the scattering, single-particle, high energy gamma with present technology in such a way that the actual efficiency of the postselection can be determined. By taking advantage of the fact that the flavor is a two-level system, we should be able to successfully construct a binary measurement. If this can be achieved, we may expect the high efficiency we wished for.

Chapter 6

Conclusion and Discussions

Weak values are quantities that have been studied from various perspectives as promising candidates directly related to the nature of quantum physics. In this dissertation, we have examined the role of the weak values focusing on its application to precision measurement. Specifically, we have considered the effectiveness of the WVA by examining two eminent experiments SHEL and USBD using photonic systems. We also explored the possibility of extending the field application by studying the CP violation of the B meson system as an example.

Our analysis of the effectiveness of the WVA is primarily based on an improved version of Lee and Tsutsui's scheme [36], where different types of uncertainties are clearly separated and studied independently. Our examination of the SHEL and USBD experiments confirmed that the WVA was indeed effective for the two experiments. In our analysis of these two experiments, we have shown that, while the statistical uncertainty can be suppressed sufficiently, the approximation uncertainty tends to dominate when the weak value is made large. We also have seen that, besides these two types of errors, what we called the intractable error becomes important as well as it may become more significant in actual experiments when we carry out the procedure of the WVA.

In this regard, we mention that since there is some ambiguity in recognizing the intractable error, it may be possible to reduce the error by specifying the mechanism of the error or by adding reasonable assumptions. In fact, if we just assume that the errors for each measurement are not correlated with each other, they will be inversely proportional to the square root of the number of particles, as a result, the WVA may increase the uncertainty caused by the postselection. *J. Sinclair et al.* have pointed out that a certain technical advantage can be gained when there exists some correlation in noise [73]. On the other hand, it has also been pointed out that the statistical and approximation uncertainties cannot be expected to be reduced by the WVA [29–31]. Our results essentially agree with their results except for the case that the (normalized) weak value is close to 1.

For both the SHEL and USBD experiments used in the analysis, we found that the parameter settings of their experiments were very close to the optimal values in terms of the relative uncertainty. Although the reason for their choice of the parameter settings is unclear, the coincidence may indicate the reliability of the present scheme of analysis we have adopted. However, our scheme is not quite sufficiently enough to account for the case $\text{Im}[A_w] = 15.9$ in the USBD experiment where the error becomes rather large.

Based on the above results, we recognize that there exist several different causes which are responsible for the errors in the actual experiments, some are reducible by using the WVA while others are not. Comparing the results with the WVA and those without it, we may determine the causes of the errors by considering the characteristics of the two

types of the measurement. This will allow us to find when the WVA can be useful as a tool for reducing the systematic error of the precision measurement.

To extend the scope of the WVA, we have examined the possibility of the WVA in the study of the CP violation of the B meson system. We have found three main results. First, the state selection of the particles created during the decay of the B meson can be used as a postselection of the state of the B meson at the time of decay. Second, the postselection at the time of decay changes the decay time distribution. In fact, the imaginary part of the weak value of the Hamiltonian, which is related to the lifetime, is seen to be extended. As a result, the effective lifetime of the B meson is prolonged by 2.6 times compared to the original lifetime. Third, our Monte-Carlo simulation shows that the accuracy of the estimation of the CP violating parameter may be improved by 20% if the final number of statistics can be ensured to be the same regardless of the postselected state (5.38) one chooses.

To perform the state selection, we need to specify the basis of the state space of the target system in a proper way and perform a measurement using the basis. To this end, the most important issue is to find an experimental means to realize the basis. One of the reasons why the optical systems are used for the WVA experiments is that such a means has already been well established. We have shown that even if it is difficult to select the state of the particles before the decay, the postselection is possible through the state selection of the particles produced by the decay. On the other hand, our results also show that insightful postselection is required for all particles involved in the decay when the decay involves multiple particles with entanglement. This fact may be an essential cause of the difficulty of the postselection in high energy experiments.

The application of the weak value to decaying particles is an interesting subject not only for the application of the WVA but also for the interest in the foundations of quantum theory. In particular, although the meter system is not used in our discussion, the same result can also be obtained by using the meter system if the assumption of the invariant mass is allowed. This may be related to the question of why time is a parameter while energy is an observable in quantum theory, and may give some indications for their fuller understanding in the future.

It would also be worthwhile to investigate some relations between the WVA and the unstable properties of the meson systems, CP violation, and the properties derived from the CPT invariance. The Hamiltonian of the mesons we used is not a Hermitian operator due to the decaying nature of the particles. Besides, the eigenvectors are not orthogonal so that the Hamiltonian cannot be diagonalized [82, 83]. This may bring an extension of the argument for more general cases described by non-Hermitian Hamiltonians for the possible role of the weak value and weak measurement.

In our discussion, we have assumed a simple relationship between the state of the B meson and the state after the decay of the B meson. According to the quantum field theory, the interaction Hamiltonian characterizes these decay phenomena, and time evolution is calculated using perturbation theory. There may still be a room for careful examination of our argument from the viewpoint of the perturbation theory.

Although our results do not necessarily show a significant improvement in accuracy when referring to the decay parameters of B meson, we have found some important results for extending the possibility of applications of the WVA. Also, the consideration of how to deal with the state before and after the decay gives a new insight for the state selection. The results are expected to be important to find the concrete method for applying the postselection to the systems in which the postselection has not been implemented so far. In addition, our theory is applicable not only to the B meson but also to other systems with unstable properties. Obviously, there should be many possibilities of applications

awaiting further investigations.

Our results will serve as a basis of expanding the applicability of the WVA and at the same time making our understandings of the foundations of the WVA more robust. We hope that the application of the WVA will expand further with the results presented in this dissertation.

Acknowledgements

I want to express my special appreciation and thanks to my supervisor and collaborator, Professor I. Tsutsui, for his valuable suggestions, discussions, and various support.

This dissertation is based on two studies I have undertaken, and I would like to thank the collaborators in each study here. The study on the effects of the WVA began with a proposal by Dr. J. Lee, and he fully supported my research. Although Mr. N. Morisawa is not a collaborator, I would also like to thank him for his early work [72] involving this study. The second half of this dissertation, the study of the new application of the WVA to B meson, would not begin without Dr. Y. Takubo's enthusiasm for applying weak measurements to high-energy experiments, and it would not have completed this study without collaborating with Dr. S. Higashino. Professor T. Higuchi and Professor A. Ishikawa also have participated in this study, and their knowledge and skills greatly have enriched our study.

Moreover, I appreciate the current members and OBs of Tsutsui Group, K. Matsuhisa, Y. Morimoto, S. Hanashiro, M. Hayakawa, K. Takeuchi, and K. Watanabe, for their support and friendship. I would also like to thank my fellows, S. Fukuchi, M. Hirasawa, T. Aoki, H. Ohta, D. Ueda, T. Ishikawa, and A. Matsumoto for the same reasons. Besides, I would like to thank the many researchers who made various discussions and comments regarding my research. I would particularly like to thank Professor Y. Shikano for his kindness and fruitful comments on the draft of this dissertation.

Finally, I owe much to my parents. Without their support and understanding, I would not have completed this work.

Appendix A

On the state of the photon and time evolution

In the text we have used the unitary operators $\hat{V}(z)$ and \hat{V}_L to describe the effect of beam propagation and its refraction at the lens. These operators are constructed with the help of the Fourier optics [71] within the so-called paraxial approximation as explained below. Since the Fourier optics is classical, we first discuss the quantum-classical correspondence of the electromagnetic field.

A.1 Quantum-Classical Correspondence

Recall that in quantum electrodynamics a field operator A_μ and E_μ are

$$\hat{A}_\mu(x) = \int \frac{d^3p}{(2\pi)^3 \sqrt{2E_{\mathbf{p}}}} \sum_{r=0}^3 (a_{\mathbf{p}}^r \epsilon_\mu^r(p) e^{-ip \cdot x} + a_{\mathbf{p}}^{r\dagger} \epsilon_\mu^{r*}(p) e^{ip \cdot x}), \quad (\text{A.1})$$

$$\begin{aligned} \hat{E}_\mu(x) &= \partial_0 \hat{A}_\mu(x) - \partial_\mu \hat{A}_0(x), \\ &= i \int \frac{d^3p}{(2\pi)^3 \sqrt{2}} \sqrt{E_{\mathbf{p}}} \sum_{r=0}^3 (a_{\mathbf{p}}^r \epsilon_\mu^r(p) e^{-ip \cdot x} - a_{\mathbf{p}}^{r\dagger} \epsilon_\mu^{r*}(p) e^{ip \cdot x}), \end{aligned} \quad (\text{A.2})$$

where ϵ_0^r is fixed to zero in this case. Regardless of whether \hat{A} or \hat{E} , and which component of those fields is selected, they are all in the format described below in common.

$$\hat{\phi}(x) = \int d^3p \sum_{r=0}^3 \left(\tilde{\phi}(|\mathbf{p}|) a_{\mathbf{p}}^r \epsilon_\mu^r e^{-ip \cdot x} + \tilde{\phi}^*(|\mathbf{p}|) a_{\mathbf{p}}^{r\dagger} \epsilon_\mu^{r*} e^{ip \cdot x} \right) \quad (\text{A.3})$$

The laser beams used in the SHEL and USBD experiments have different wavelengths, but both are monochromatic light so that the above $\tilde{\phi}(|\mathbf{p}|)$ is a constant in their cases. To obtain the corresponding classical field, we evaluate the following amplitude.

$$\phi(x) = \langle 0 | \hat{\phi}(x) | \Psi \rangle \quad (\text{A.4})$$

$$= \phi_0 \sum_{s'=1,2} \int d^3p \Psi(s', \mathbf{p}) \epsilon^{s'} e^{-ip \cdot x} \quad (\text{A.5})$$

It follows that, since $\phi(x)$ in (A.6) satisfies the field equation (which is fulfilled by $\hat{\phi}(x)$), so does the wave function $\gamma(\mathbf{x}, t)$.

$$\phi(x) := \langle 0 | \hat{\phi}(x) | \gamma \rangle \propto \int d^3k e^{-i\mathbf{k} \cdot \mathbf{x} + i\omega t} \gamma(\mathbf{k}, t). \quad (\text{A.6})$$

A.2 Free propagation

We introduce the coordinate in which the beam propagates along the z direction and the shift obtained by the interaction (3.2) takes place in the x direction. The beam is approximated by the plane wave in the z direction with momentum p_z with a tiny spread of the p_x component in the x direction, which are related to the respective wave numbers by $p_z = \hbar k_z$ and $p_x = \hbar k_x$. Assuming that the angular frequency ω of the beam is fixed, the profile of the electromagnetic field can be written as

$$\phi(\mathbf{x}, t) = g(\mathbf{x})e^{i(k_0z - \omega t)}. \quad (\text{A.7})$$

In this form, the wave equation reads

$$\begin{aligned} 0 &= \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi(\mathbf{x}, t) \\ &= e^{i(k_0z - \omega t)} \left(-\nabla^2 - 2ik_0 \frac{\partial}{\partial z} \right) g(\mathbf{x}). \end{aligned} \quad (\text{A.8})$$

The paraxial approximation is valid when the conditions

$$\left| \frac{\partial^2 g(\mathbf{x})}{\partial z^2} \right| \ll \left| \frac{\partial^2 g(\mathbf{x})}{\partial x^2} \right|, \left| \frac{\partial^2 g(\mathbf{x})}{\partial y^2} \right|, \left| k_0 \frac{\partial g(\mathbf{x})}{\partial z} \right|, \quad (\text{A.9})$$

are fulfilled. To check the validity of these, we consider the Fourier expansion of $\phi(\mathbf{x}, t)$,

$$\phi(\mathbf{x}, t) = \int d^3k \tilde{\phi}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}. \quad (\text{A.10})$$

Plugging (A.10) into (A.8), we obtain

$$k_x^2 + k_y^2 + k_z^2 = k_0^2, \quad (\text{A.11})$$

for each mode \mathbf{k} . Combining (A.7) with (A.10), one finds

$$g(\mathbf{x}) = \int d^3k \tilde{\phi}(\mathbf{k}) e^{i(k_x x + k_y y + (k_z - k_0)z)}. \quad (\text{A.12})$$

Then the conditions (A.9) are satisfied if

$$|k_z - k_0| \ll |k_x|, |k_y|, \sqrt{|k_0(k_z - k_0)|}, \quad (\text{A.13})$$

which are assured if we just have

$$k_x, k_y \ll k_0 \quad (\text{A.14})$$

on account of (A.11). The beam used in the experiments which we analyzed in the paper indeed fulfills these conditions (A.14).

Within the paraxial approximation (A.9), we therefore obtain from the wave equation (A.8),

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2ik_0 \frac{\partial}{\partial z} \right) g(\mathbf{x}) = 0. \quad (\text{A.15})$$

Upon using (A.12), the equation (A.15) can be solved as

$$g(\mathbf{x}) = \int d^3k \tilde{\phi}(\mathbf{k}) e^{i(k_x x + k_y y - \frac{1}{2k_0}(k_x^2 + k_y^2)z)}, \quad (\text{A.16})$$

which implies

$$\phi(\mathbf{x}, t) = \int d^3k \tilde{\phi}(\mathbf{k}) e^{i(k_x x + k_y y - \frac{1}{2k_0}(k_x^2 + k_y^2)z)} e^{i(k_0 z - \omega t)}. \quad (\text{A.17})$$

Now, in view of (A.5), one may introduce the unitary operator representing the propagation effect,

$$V_P = e^{-i \frac{k_x^2 + k_y^2}{2k_0} l}, \quad (\text{A.18})$$

and thereby express the relation between $\phi(x, y, 0, t)$ to $\phi(x, y, l, t)$ in terms of the wave function $\gamma(\mathbf{x}, t)$ as

$$\gamma(x, y, l, t) = \langle x, y, l | \gamma \rangle = \langle x, y, 0 | V_P | \gamma \rangle. \quad (\text{A.19})$$

The unitary operators (3.73) used in the text arise when we have the shift only in the x direction. The state appearing above corresponds to the meter state $|\xi\rangle$ there.

A.3 Lens as a unitary operator

In addition to the propagation effect, we also need to take into account the effect of the lens used in the experiment. Since the profile function (A.17) of the laser beam is a superposition of various plane waves with fixed \mathbf{k} , it is enough to consider the effect only on the plane wave.

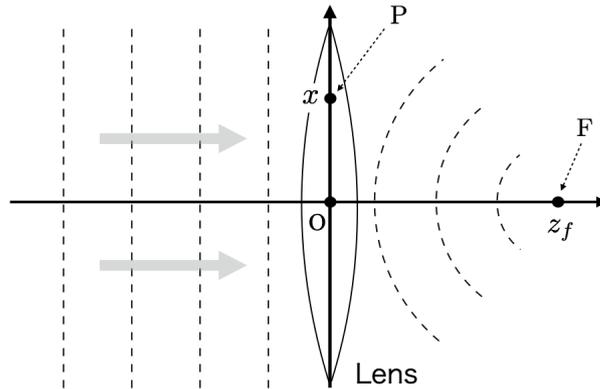


Figure A.1: The lens and the wave surface after passing through the lens

For this, we first observe that, during the passage of the plane wave from the point P in the lens to the focal point F (see Figure A.1), the plane wave acquires the extra phase,

$$\vartheta(x) = k_0 \left(\sqrt{z_f^2 + x^2} - z_f \right), \quad (\text{A.20})$$

compared to the passage in the center from O to F, where z_f is the focal length of the lens. If the lens is thin enough so that the thickness effect can be ignored, then for $z_f \gg x$ we have

$$\vartheta(x) \simeq k_0 \frac{x^2}{2z_f}, \quad (\text{A.21})$$

which yields the phase change $e^{-ik_0 \frac{x^2}{2z_f}}$ for each of the plane waves. As a result, the effect of the lens can be incorporated if we modify the profile function (A.17), or equivalently the wave function $\gamma(\mathbf{x}, t)$, by inserting the unitary operator representing the lens effect,

$$V_L = e^{-i \frac{k_0}{2z_f} \hat{x}^2}, \quad (\text{A.22})$$

at an appropriate position as we have done in the formula (A.19) in the case of the propagation effect.

Appendix B

Some useful results

Some of the arguments and calculations used in the text are presented here in detail.

B.1 Momentum as the generator of translation

It is known that the momentum \hat{P} is the generator of the translation [84, 85]. In fact, we find

$$\begin{aligned} e^{-i\theta\hat{P}} |\chi\rangle &= e^{-i\theta\hat{P}} \int_{-\infty}^{\infty} dx \chi(x) |x\rangle \\ &= e^{-i\theta\hat{P}} \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx \chi(x) |p\rangle \langle p|x\rangle \\ &= e^{-i\theta\hat{P}} \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx \chi(x) \frac{e^{-ipx}}{\sqrt{2\pi}} |p\rangle \\ &= \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx \chi(x) \frac{e^{-ip(x+\theta)}}{\sqrt{2\pi}} |p\rangle \\ &= \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx \chi(x) \langle p|x+\theta\rangle |p\rangle \\ &= \int_{-\infty}^{\infty} dx \chi(x-\theta) |x\rangle. \end{aligned} \tag{B.1}$$

This means that the initial wave function $\chi(x)$ is transformed to $\chi(x-\theta)$.

B.2 The relation between the eigenstates of the flavor and the Hamiltonian

This section is based on section 2.1 of the paper [83]. This paper deals with CP violation in kaons, but it is known that it can be basically used as a theory for B mesons by reading $|K^0\rangle$ as $|B^0\rangle$ and $|\bar{K}^0\rangle$ as $|\bar{B}^0\rangle$ [57, Secs. 69, 74].

First, we define discrete symmetry C , P and T (charge, parity conjugation and time reversal) as quantum-mechanical operators that act on flavor eigenstates in the following way.

$$\begin{aligned} \mathcal{C} |B^0\rangle &= e^{i\alpha_c} |\bar{B}^0\rangle, & \mathcal{C} |\bar{B}^0\rangle &= e^{-i\alpha_c} |B^0\rangle, \\ \mathcal{P} |B^0\rangle &= -|B^0\rangle, & \mathcal{P} |\bar{B}^0\rangle &= -|\bar{B}^0\rangle, \\ \mathcal{T} |B^0\rangle &= e^{i(\theta-\alpha_c)} |B^0\rangle, & \mathcal{T} |\bar{B}^0\rangle &= e^{i(\theta+\alpha_c)} |\bar{B}^0\rangle, \end{aligned} \tag{B.2}$$

where α_c is arbitrary phase parameters.

\mathcal{T} is an antilinear operator. An operator \mathcal{T} is said to be antilinear if

$$\mathcal{T}(c_1 |\psi_1\rangle + c_2 |\psi_2\rangle) = c_1^* \mathcal{T} |\psi_1\rangle + c_2^* \mathcal{T} |\psi_2\rangle \quad (\text{B.3})$$

for all $c_1, c_2 \in \mathbb{C}$ and all states $|\psi_1\rangle, |\psi_2\rangle$.

We assume the Hamiltonian has the CPT invariance. Considering the antilinearity of \mathcal{T} operator, the CPT invariance for the non-Hermitian effective Hamiltonian \hat{H} should be imposed by [86]

$$CPT \hat{H} (CPT)^{-1} = \hat{H}^\dagger. \quad (\text{B.4})$$

When the components of the Hamiltonian are defined by

$$\hat{H} = h_{11} |B^0\rangle \langle B^0| + h_{12} |\bar{B}^0\rangle \langle B^0| + h_{21} |B^0\rangle \langle \bar{B}^0| + h_{22} |\bar{B}^0\rangle \langle \bar{B}^0|, \quad (\text{B.5})$$

we can evaluate the left-hand side of (B.4) specifically.

$$CPT \hat{H} (CPT)^{-1} = h_{11}^* |\bar{B}^0\rangle \langle \bar{B}^0| + h_{12}^* |B^0\rangle \langle \bar{B}^0| + h_{21}^* |\bar{B}^0\rangle \langle B^0| + h_{22}^* |B^0\rangle \langle B^0| \quad (\text{B.6})$$

As a result of the calculation, the condition

$$h_{11} = h_{22} =: h \quad (\text{B.7})$$

should be satisfied. Then, the eigenvalues are

$$h \pm \sqrt{h_{12} h_{21}}, \quad (\text{B.8})$$

and the eigenvectors are respectively

$$p |B^0\rangle \pm q |\bar{B}^0\rangle, \quad (\text{B.9})$$

where the coefficients p, q are satisfying

$$\frac{p}{q} = \sqrt{\frac{h_{12}}{h_{21}}}. \quad (\text{B.10})$$

B.3 Monotonicity of the Lower-bound Function

We prove that the lower-bound function $r(\delta_{\text{st}}; N, q)$ given in (4.32) is a monotonically increasing function¹ with respect to all of its parameters δ_{st} , N and q .

With the binomial distribution (4.33), we first observe that, as a finite weighted average of monotonically increasing functions (4.27), monotonicity with respect to δ_{st} is trivial. To see the monotonicity with respect to N , for convenience we extend formally the range of n in $\binom{N}{n}$ from non-negative integers to integers $n \in \mathbb{Z}$ by defining $\binom{N}{n} = 0$ for negative n and thereby confirm the formula,

$$\text{Bi}(n; N+1, q) - \text{Bi}(n; N, q) = q (\text{Bi}(n-1; N, q) - \text{Bi}(n; N, q)), \quad (\text{B.11})$$

¹A function f is said to be monotonically increasing if $a < b$ implies $f(a) \leq f(b)$. Specifically, f is called strictly monotonically increasing if $a < b$ implies $f(a) < f(b)$.

for $N \in \mathbb{N}$, which can be directly obtained by a simple application of the recursive formula, $\binom{N+1}{n} = \binom{N}{n} + \binom{N}{n-1}$, valid for $N > 0$. This allows us to rewrite

$$\begin{aligned}
r(\delta_{\text{st}}; N+1, q) - r(\delta_{\text{st}}; N, q) &= \sum_{n=1}^{N+1} T(\delta_{\text{st}}; n) \left(\text{Bi}(n; N+1, q) - \text{Bi}(n; N, q) \right) \\
&= q \sum_{n=1}^{N+1} T(\delta_{\text{st}}; n) \left(\text{Bi}(n-1; N, q) - \text{Bi}(n; N, q) \right) \\
&= q \sum_{n=1}^N \left(T(\delta_{\text{st}}; n+1) - T(\delta_{\text{st}}; n) \right) \text{Bi}(n; N, q) + qT(\delta_{\text{st}}; 1) \text{Bi}(0; N, q),
\end{aligned} \tag{B.12}$$

which is always non-negative due to the monotonicity of the function (4.27) with respect to n . Finally, as for q , first note that the monotonicity trivially holds for $N = 0$, since $r(\delta_{\text{st}}; 0, q) = 0$ is a constant function. For $N > 0$, we may utilize the formula,

$$\frac{\partial \text{Bi}(n; N, q)}{\partial q} = N \left(\text{Bi}(n-1; N-1, q) - \text{Bi}(n; N-1, q) \right) \tag{B.13}$$

which can be readily obtained from $n \binom{N}{n} = N \binom{N-1}{n-1}$ valid for $N > 0$. We then have

$$\begin{aligned}
\frac{\partial r(\delta_{\text{st}}; N, q)}{\partial q} &= \sum_{n=1}^N T(\delta_{\text{st}}; n) \frac{\partial \text{Bi}(n; N, q)}{\partial q} \\
&= N \sum_{n=1}^N T(\delta_{\text{st}}; n) \left(\text{Bi}(n-1; N-1, q) - \text{Bi}(n; N-1, q) \right) \\
&= N \sum_{n=1}^N \left(T(\delta_{\text{st}}; n+1) - T(\delta_{\text{st}}; n) \right) \text{Bi}(n; N-1, q) + N \cdot T(\delta_{\text{st}}; 1) \text{Bi}(0; N-1, q)
\end{aligned} \tag{B.14}$$

which is always non-negative due to the monotonicity of the function (4.27) with respect to n : this completes the proof of the desired statement.

B.4 Uncorrelated intractable error

To see the intractable error for each measurement clearly, we define the intractable error for i th measurement as

$$\delta x_i := \tilde{X}_i - X_i. \tag{B.15}$$

For the average of n times measurements, the expected value of the difference between \tilde{X}_n , which is calculated with using the actual measurement results and X_n , which is calculated with using the true values is

$$\sqrt{\mathbf{E} \left(\left| \tilde{X}_n - X_n \right|^2 \right)} = \sqrt{\mathbf{E} \left(\left| \frac{1}{n} \sum_{i=1}^n \delta x_i \right|^2 \right)} = \frac{1}{n} \sqrt{\mathbf{E} \left(\sum_{i=1}^n \delta x_i^2 + \sum_{i \neq j} \delta x_i \delta x_j \right)} \tag{B.16}$$

We can use the linearity of the expectation value and the non-correlation of the error δx_i so that we find

$$(B.16) = \frac{1}{n} \sqrt{\mathbf{E} \left(\sum_{i=1}^n \delta x_i^2 \right) + \mathbf{E} \left(\sum_{i \neq j} \delta x_i \delta x_j \right)} = \frac{1}{n} \sqrt{\mathbf{E} \left(\sum_{i=1}^n \delta x_i^2 \right)}. \quad (B.17)$$

Finally, we can obtain the inequality

$$(B.17) \leq \frac{1}{n} \sqrt{n \delta_{\text{int}}^2} = \frac{\delta_{\text{int}}}{\sqrt{n}}. \quad (B.18)$$

This is equal to the inequality (4.52).

B.5 Calculation involving the Gaussian integrals

In the calculations in Chapter 4, we use the Gaussian integrals many times. In this chapter, the formulas required for the calculations dealt with in Chapter 4 and their derivations are described.

$$I_1(\alpha) := \int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}} \quad (B.19)$$

$$\int_{-\infty}^{\infty} dx x^2 e^{-\alpha x^2} = \frac{1}{2\alpha} \sqrt{\frac{\pi}{\alpha}} \quad \left(= -\frac{d}{d\alpha} I_1(\alpha) \right) \quad (B.20)$$

$$I_2(\alpha, \beta) := \int_{-\infty}^{\infty} dx \cos(\beta x) e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}} \quad (B.21)$$

$$\int_{-\infty}^{\infty} dx x \sin(\beta x) e^{-\alpha x^2} = \frac{\beta}{2\alpha} \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}} \quad \left(= -\frac{d}{d\beta} I_2(\alpha, \beta) \right) \quad (B.22)$$

$$\int_{-\infty}^{\infty} dx x^2 \cos(\beta x) e^{-\alpha x^2} = \left(\frac{1}{2\alpha} - \frac{\beta^2}{4\alpha^2} \right) \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}} \quad \left(= -\frac{d}{d\alpha} I_2(\alpha, \beta) \right) \quad (B.23)$$

In this case, each part that appears in the calculation can be obtained as follows. In (4.23), $|\zeta\rangle$ is given by

$$|\zeta\rangle = \int dx (\cos(\theta x) - i A_w \sin(\theta x) e^{-\alpha' x^2} |x\rangle, \quad (B.24)$$

where α' is

$$\alpha' = \frac{k_0(2\alpha l_f + ik_0)}{2k_0(l_f - z) + 4i\alpha z l_f} \quad (B.25)$$

$$= \frac{2k_0^2 \alpha l_f^2 + ik_0 \{k_0^2(l_f - z) - 4\alpha^2 z l_f^2\}}{2k_0^2(l_f - z)^2 + 8\alpha^2 z^2 l_f^2} \quad (B.26)$$

Also, $|\zeta_P\rangle$ is given by

$$|\zeta_P\rangle = \int dp (\cos(\theta p) - i A_w \sin(\theta p) e^{-\alpha' p^2} |p\rangle, \quad (B.27)$$

where α'_P is

$$\alpha'_P = i \frac{z}{2k_0} + \frac{1}{4\alpha + i \frac{2k_0}{l_f}} \quad (\text{B.28})$$

$$= \frac{2k_0\alpha l_f^2 + i(4\alpha^2 l_f^2 z - k_0^2(l_f - z))}{2k_0(4\alpha^2 l_f^2 + k_0^2)} = \left(\frac{1}{4\alpha'} \right) \quad (\text{B.29})$$

The only substantial difference between (B.24) and (B.27) is only the selection of basis and it does not make differences in most of the following calculations. Therefore, the part calculated in the case of using (B.24) below and affects the result.

$$\begin{aligned} \langle \zeta | \zeta \rangle &= \int dx |\cos(\theta x) - iA_w \sin(\theta x)|^2 e^{-2\text{Re}[\alpha']x^2} \\ &= \int dx \{ \cos^2(\theta x) + |A_w|^2 \sin^2(\theta x) + \text{Im}[A_w] \sin(2\theta x) \} e^{-2\text{Re}[\alpha']x^2} \\ &= \frac{1 + |A_w|^2}{2} \int dx e^{-2\text{Re}[\alpha']x^2} + \frac{1 - |A_w|^2}{2} \int dx e^{-2\text{Re}[\alpha']x^2} \cos(2\theta x) \\ &= \frac{1 + |A_w|^2}{2} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} + \frac{1 - |A_w|^2}{2} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \\ &= \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \right) \end{aligned} \quad (\text{B.30})$$

If we calculate $\langle \zeta_P | \zeta_P \rangle$, the variable α' in (B.30) should be changed to α'_P .

$$\begin{aligned} \langle \zeta | \hat{X} | \zeta \rangle &= \int dx x \cos^2(\theta x) e^{-2\text{Re}[\alpha']x^2} + |A_w|^2 \int dx x \sin^2(\theta x) e^{-2\text{Re}[\alpha']x^2} \\ &\quad + \text{Im}[A_w] \int dx x \sin(2\theta x) e^{-2\text{Re}[\alpha']x^2} \\ &= \text{Im}[A_w] \frac{\theta}{2\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \end{aligned} \quad (\text{B.31})$$

$\langle \zeta_P | \hat{P} | \zeta_P \rangle$ can be obtained by changing the variable α' in (B.31) to α'_P .

$$\langle \zeta_P | \hat{P} | \zeta_P \rangle = \text{Im}[A_w] \frac{\theta}{2\text{Re}[\alpha'_P]} \sqrt{\frac{\pi}{2\text{Re}[\alpha'_P]}} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \quad (\text{B.32})$$

$$\begin{aligned} \langle \zeta | \hat{X}^2 | \zeta \rangle &= \int dx x^2 \cos^2(\theta x) e^{-2\text{Re}[\alpha']x^2} + |A_w|^2 \int dx x^2 \sin^2(\theta x) e^{-2\text{Re}[\alpha']x^2} \\ &\quad + \text{Im}[A_w] \int dx x^2 \sin(2\theta x) e^{-2\text{Re}[\alpha']x^2} \\ &= \frac{1 + |A_w|^2}{2} \int dx x^2 e^{-2\text{Re}[\alpha']x^2} + \frac{1 - |A_w|^2}{2} \int dx x^2 e^{-2\text{Re}[\alpha']x^2} \cos(2\theta x) \\ &= \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \frac{1}{4\text{Re}[\alpha']} \left(\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(1 - \frac{\theta^2}{\text{Re}[\alpha']} \right) \right) \end{aligned} \quad (\text{B.33})$$

In our calculation, we need the expectation value of the momentum \hat{P} or \hat{P}^2 .

$$\begin{aligned} \hat{P} | \zeta \rangle &= -i \int dx \frac{d}{dx'} \left[(\cos(\theta x') - iA_w \sin(\theta x')) e^{-\alpha' x'^2} \right]_{x'=x} | x \rangle. \\ &= -i \int dx \left\{ (-\theta \sin(\theta x) - iA_w \theta \cos(\theta x)) e^{-\alpha' x^2} - 2(\cos(\theta x) - iA_w \sin(\theta x)) (\alpha' x e^{-\alpha' x^2}) \right\} | x \rangle. \end{aligned} \quad (\text{B.34})$$

Also,

$$\begin{aligned}
\hat{P}^2|\zeta\rangle &= - \int dx \frac{d}{dx'} \left\{ (-\theta \sin(\theta x') - iA_w \theta \cos(\theta x')) e^{-\alpha' x'^2} \right. \\
&\quad \left. + (\cos(\theta x') - iA_w \sin(\theta x')) (-2\alpha' x' e^{-\alpha' x'^2}) \right\}_{x'=x} |x\rangle \\
&= - \int dx \left\{ (-\theta^2 \cos(\theta x) + iA_w \theta^2 \sin(\theta x)) e^{-\alpha' x^2} \right. \\
&\quad \left. + 2(-\theta \sin(\theta x) - iA_w \theta \cos(\theta x)) (-2\alpha' x e^{-\alpha' x^2}) \right. \\
&\quad \left. + (\cos(\theta x) - iA_w \sin(\theta x)) (-2\alpha' e^{-\alpha' x^2} + 4\alpha'^2 x^2 e^{-\alpha' x^2}) \right\} |x\rangle
\end{aligned} \tag{B.35}$$

so we can obtain

$$\begin{aligned}
\langle \zeta | \hat{P} | \zeta \rangle &= -i \int dx \left\{ (-iA_w^* \theta \sin^2(\theta x) - iA_w \theta \cos^2(\theta x)) e^{-2\text{Re}[\alpha'] x^2} \right. \\
&\quad \left. + (iA_w^* \sin(\theta x) \cos(\theta x) - iA_w \sin(\theta x) \cos(\theta x)) (-2\alpha' x e^{-2\text{Re}[\alpha'] x^2}) \right\} \\
&= -i \int dx \left\{ (-i\theta \text{Re}[A_w] + \theta \text{Im}[A_w] \cos(2\theta x)) e^{-2\text{Re}[\alpha'] x^2} - 2\text{Im}[A_w] \sin(2\theta x) (\alpha' x e^{-2\text{Re}[\alpha'] x^2}) \right\} \\
&= -\theta \text{Re}[A_w] \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} - i\theta \text{Im}[A_w] \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} + i\alpha' \text{Im}[A_w] \frac{\theta}{\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \\
&= -\theta \text{Re}[A_w] \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} - \theta \text{Im}[A_w] \frac{\text{Im}[\alpha']}{\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \\
&= -\theta \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\text{Re}[A_w] + \text{Im}[A_w] \frac{\text{Im}[\alpha']}{\text{Re}[\alpha']} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \right)
\end{aligned} \tag{B.36}$$

Since the position operator \hat{X} on momentum basis can be expressed as $\hat{X} = i\frac{\partial}{\partial p}$, the expected value of the position with respect to the state $|\zeta_P\rangle$ can be expressed as follows.

$$\langle \zeta_P | \hat{X} | \zeta_P \rangle = \theta \sqrt{\frac{\pi}{2\text{Re}[\alpha'_P]}} \left(\text{Re}[A_w] + \text{Im}[A_w] \frac{\text{Im}[\alpha'_P]}{\text{Re}[\alpha'_P]} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \right) \tag{B.37}$$

Also, we need the expectation value of the squared momentum. This is

$$\begin{aligned}
\langle \zeta | \hat{P}^2 | \zeta \rangle &= - \int dx \left\{ (-\theta^2 \cos^2(\theta x) - |A_w|^2 \theta^2 \sin^2(\theta x)) e^{-2\text{Re}[\alpha'] x^2} \right. \\
&\quad \left. + 2(-\theta \sin(\theta x) \cos(\theta x) + |A_w|^2 \theta \sin(\theta x) \cos(\theta x)) (-2\alpha' x e^{-2\text{Re}[\alpha'] x^2}) \right. \\
&\quad \left. + (\cos^2(\theta x) + |A_w|^2 \sin^2(\theta x)) (-2\alpha' e^{-2\text{Re}[\alpha'] x^2} + 4\alpha'^2 x^2 e^{-2\text{Re}[\alpha'] x^2}) \right\}, \\
&= - \int dx \left\{ (-\theta^2 - 2\alpha') \frac{1 + |A_w|^2}{2} e^{-2\text{Re}[\alpha'] x^2} - 2\alpha' \theta (-1 + |A_w|^2) \sin(2\theta x) x e^{-2\text{Re}[\alpha'] x^2} \right. \\
&\quad \left. + (-\theta^2 - 2\alpha') \frac{1 - |A_w|^2}{2} \cos(2\theta x) e^{-2\text{Re}[\alpha'] x^2} + 4\alpha'^2 \frac{1 + |A_w|^2}{2} x^2 e^{-2\text{Re}[\alpha'] x^2} \right. \\
&\quad \left. + 4\alpha'^2 \frac{1 - |A_w|^2}{2} x^2 \cos(2\theta x) e^{-2\text{Re}[\alpha'] x^2} \right\},
\end{aligned} \tag{B.38}$$

$$\begin{aligned}
&= (\theta^2 + 2\alpha') \frac{1 + |A_w|^2}{2} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} + 2\alpha'\theta(-1 + |A_w|^2) \frac{\theta}{2\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \\
&\quad + (\theta^2 + 2\alpha') \frac{1 - |A_w|^2}{2} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} - 4\alpha'^2 \frac{1 + |A_w|^2}{2} \frac{1}{4\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \\
&\quad - 4\alpha'^2 \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\frac{1}{4\text{Re}[\alpha']} - \frac{\theta^2}{4\text{Re}[\alpha']^2} \right), \tag{B.39}
\end{aligned}$$

$$\begin{aligned}
&= \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\left(\theta^2 + \text{Re}[\alpha'] + \frac{\text{Im}[\alpha']^2}{\text{Re}[\alpha']} \right) \frac{1 + |A_w|^2}{2} \right. \\
&\quad \left. + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(\text{Re}[\alpha'] + \frac{\text{Im}[\alpha']^2}{\text{Re}[\alpha']} - \theta^2 \frac{\text{Im}[\alpha']^2}{\text{Re}[\alpha']^2} \right) \right). \tag{B.40}
\end{aligned}$$

In the same way, we can obtain

$$\begin{aligned}
\langle \zeta_P | \hat{X}^2 | \zeta_P \rangle &= \sqrt{\frac{\pi}{2\text{Re}[\alpha'_P]}} \left(\left(\theta^2 + \text{Re}[\alpha'_P] + \frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]} \right) \frac{1 + |A_w|^2}{2} \right. \\
&\quad \left. + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \left(\text{Re}[\alpha'_P] + \frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]} - \theta^2 \frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]^2} \right) \right). \tag{B.41}
\end{aligned}$$

To evaluate $\langle \zeta | \{ \hat{X}, \hat{P} \} | \zeta \rangle$, we use

$$\{ \hat{X}, \hat{P} \} = 2\hat{X}\hat{P} - i. \tag{B.42}$$

Then, $\langle \zeta | \hat{X}\hat{P} | \zeta \rangle$ is

$$\begin{aligned}
\langle \zeta | \hat{X}\hat{P} | \zeta \rangle &= -i \int dx \{ (-x\theta \cos(\theta x) \sin(\theta x))(1 - |A_w|^2)e^{-2\text{Re}[\alpha']x^2} \\
&\quad - 2\alpha'x^2 |\cos(\theta x) - iA_w \sin(\theta x)|^2 e^{-2\text{Re}[\alpha']x^2} \}, \\
&= i \int dx \{ (x\theta \sin(2\theta x)) \frac{1 - |A_w|^2}{2} e^{-2\text{Re}[\alpha']x^2} \\
&\quad + 2\alpha'x^2 (\cos^2(\theta x) + |A_w|^2 \sin^2(\theta x)) e^{-2\text{Re}[\alpha']x^2} \}, \\
&= i \int dx \{ (x\theta \sin(2\theta x)) \frac{1 - |A_w|^2}{2} e^{-2\text{Re}[\alpha']x^2} + 2\alpha' \left(\frac{1 + |A_w|^2}{2} \right) x^2 e^{-2\text{Re}[\alpha']x^2} \\
&\quad + 2\alpha' \left(\frac{1 - |A_w|^2}{2} \right) x^2 e^{-2\text{Re}[\alpha']x^2} \cos(2\theta x) \}, \\
&= i\theta \frac{1 - |A_w|^2}{2} \frac{\theta}{2\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} + 2i\alpha' \left(\frac{1 + |A_w|^2}{2} \right) \frac{1}{4\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \\
&\quad + 2i\alpha' \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\frac{1}{4\text{Re}[\alpha']} - \frac{\theta^2}{4\text{Re}[\alpha']^2} \right), \\
&= \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\frac{1 + |A_w|^2}{2} \left(\frac{i}{2} - \frac{\text{Im}[\alpha']}{2\text{Re}[\alpha']} \right) + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(\frac{i\alpha'}{2\text{Re}[\alpha']} + \frac{\theta^2 \text{Im}[\alpha']}{2\text{Re}[\alpha']^2} \right) \right). \tag{B.43}
\end{aligned}$$

Combining (B.43) and (B.30), the solution is

$$\begin{aligned}
\langle \xi | \{ \hat{X}, \hat{P} \} | \zeta \rangle &= 2 \langle \zeta | \hat{X} \hat{P} | \zeta \rangle - i \langle \zeta | \zeta \rangle, \\
&= \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\frac{1 + |A_w|^2}{2} \left(i - \frac{\text{Im}[\alpha']}{\text{Re}[\alpha']} \right) + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(\frac{i\alpha'}{\text{Re}[\alpha']} + \frac{\theta^2 \text{Im}[\alpha']}{\text{Re}[\alpha']^2} \right) \right) \\
&\quad - i \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \right), \\
&= -\frac{\text{Im}[\alpha']}{\text{Re}[\alpha']} \sqrt{\frac{\pi}{2\text{Re}[\alpha']}} \left(\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(1 - \frac{\theta^2}{\text{Re}[\alpha']} \right) \right). \tag{B.44}
\end{aligned}$$

Also, its $|\zeta_P\rangle$ version is

$$\begin{aligned}
\langle \zeta_P | \{ \hat{X}, \hat{P} \} | \zeta_P \rangle &= 2 \langle \zeta_P | \hat{P} \hat{X} | \zeta_P \rangle + i \langle \zeta_P | \zeta_P \rangle, \\
&= \frac{\text{Im}[\alpha'_P]}{\text{Re}[\alpha'_P]} \sqrt{\frac{\pi}{2\text{Re}[\alpha'_P]}} \left(\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \left(1 - \frac{\theta^2}{\text{Re}[\alpha'_P]} \right) \right). \tag{B.45}
\end{aligned}$$

In $\hat{Y} = \hat{X}$ case, we obtain

$$\langle \xi | \hat{X} | \xi \rangle = \frac{\text{Im}[A_w] \frac{\theta}{2\text{Re}[\alpha']} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}{\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}, \tag{B.46}$$

$$\langle \xi | \hat{X}^2 | \xi \rangle = \frac{\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(1 - \frac{\theta^2}{\text{Re}[\alpha']} \right)}{4\text{Re}[\alpha'] \left(\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \right)}, \tag{B.47}$$

$$\langle \xi | \hat{P} | \xi \rangle = \frac{-\theta \text{Re}[A_w] - \theta \text{Im}[A_w] \frac{\text{Im}[\alpha']}{\text{Re}[\alpha']} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}{\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}, \tag{B.48}$$

$$\langle \xi | \hat{P}^2 | \xi \rangle = \frac{\left(\theta^2 + \text{Re}[\alpha'] + \frac{\text{Im}[\alpha']^2}{\text{Re}[\alpha']} \right) \frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(\text{Re}[\alpha'] + \frac{\text{Im}[\alpha']^2}{\text{Re}[\alpha']} - \theta^2 \frac{\text{Im}[\alpha']^2}{\text{Re}[\alpha']^2} \right)}{\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}, \tag{B.49}$$

$$\langle \xi | \{ \hat{X}, \hat{P} \} | \xi \rangle = -\frac{\text{Im}[\alpha']}{\text{Re}[\alpha']} \frac{\left(\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}} \left(1 - \frac{\theta^2}{\text{Re}[\alpha']} \right) \right)}{\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha']}}}. \tag{B.50}$$

Also, in $\hat{Y} = \hat{P}$ case, we obtain

$$\langle \xi | \hat{X} | \xi \rangle = \frac{\theta \text{Re}[A_w] + \theta \text{Im}[A_w] \frac{\text{Im}[\alpha'_P]}{\text{Re}[\alpha'_P]} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}{\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}, \tag{B.51}$$

$$\langle \xi | \hat{X}^2 | \xi \rangle = \frac{\left(\theta^2 + \text{Re}[\alpha'_P] + \frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]} \right) \frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \left(\text{Re}[\alpha'_P] + \frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]} - \theta^2 \frac{\text{Im}[\alpha'_P]^2}{\text{Re}[\alpha'_P]^2} \right)}{\frac{1 + |A_w|^2}{2} + \frac{1 - |A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}, \tag{B.52}$$

$$\langle \xi | \hat{P} | \xi \rangle = \frac{\text{Im}[A_w] \frac{\theta}{2\text{Re}[\alpha'_P]} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}, \quad (\text{B.53})$$

$$\langle \xi | \hat{P}^2 | \xi \rangle = \frac{\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \left(1 - \frac{\theta^2}{\text{Re}[\alpha'_P]}\right)}{4\text{Re}[\alpha'] \left(\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}\right)}, \quad (\text{B.54})$$

$$\langle \xi | \{\hat{X}, \hat{P}\} | \xi \rangle = \frac{\text{Im}[\alpha'_P] \left(\frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}} \left(1 - \frac{\theta^2}{\text{Re}[\alpha'_P]}\right)\right)}{\text{Re}[\alpha'_P] \frac{1+|A_w|^2}{2} + \frac{1-|A_w|^2}{2} e^{-\frac{\theta^2}{2\text{Re}[\alpha'_P]}}}. \quad (\text{B.55})$$

Bibliography

- [1] J. Schwinger. On Quantum-Electrodynamics and the Magnetic Moment of the Electron. *Phys. Rev.*, 73:416, 1948. doi:10.1103/PhysRev.73.416.
- [2] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. CRC-press, 1995.
- [3] E. Schrödinger. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.*, 28:1049, 1926. doi:10.1103/PhysRev.28.1049.
- [4] W. Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43:172, 1927. doi:10.1007/BF01397280.
- [5] J. von Neumann. *Die Mathematische Grundlagen der Quantenmechanik*. Springer, 1932.
- [6] H. M. Wiseman and G. J. Milburn. *Quantum Measurement and Control*. Cambridge, second edition, 2014.
- [7] Y. Aharonov, P. G. Bergmann, and L. Lebowitz. Time Symmetry in the Quantum Process of Measurement. *Phys. Rev.*, 134:B1410, 1964. doi:10.1103/PhysRev.134.B1410.
- [8] 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. doi:10.1103/PhysRevLett.60.1351.
- [9] Y. Aharonov and L. Vaidman. Complete description of a quantum system at a given time. *J. Phys. A: Math. Gen.*, 24:2315, 1991. doi:10.1088/0305-4470/24/10/018.
- [10] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg. Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer. *Science*, 332(6034):1170, 2011. doi:10.1126/science.1202218.
- [11] A. Danan, D. Farfurnik, S. Bar-Ad, and L. Vaidman. Asking Photons Where They Have Been. *Phys. Rev. Lett.*, 111:240402, 2013. arXiv:1304.7469, doi:10.1103/PhysRevLett.111.240402.
- [12] T. Mori and I. Tsutsui. Weak value and the wave-particle Duality. *Quantum Stud.: Math. Found.*, 2(4):371, 2015. arXiv:1410.0787, doi:10.1007/s40509-015-0039-5.
- [13] L. Vaidman and I. Tsutsui. When photons are lying about where they have been. *Entropy*, 20:538, 2018. arXiv:1806.00695, doi:10.3390/e20070538.

- [14] 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. doi:10.1103/PhysRevLett.66.1107.
- [15] D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann. Fast light, slow light, and phase singularities: A connection to generalized weak values. *Phys. Rev. Lett.*, 92:043601, 2004. arXiv:quant-ph/0310048, doi:10.1103/PhysRevLett.92.043601.
- [16] K.J. Resch, J.S. Lundeen, and A.M. Steinberg. Experimental realization of the quantum box problem. *Phys. Lett. A*, 324(2):125–131, 2004. arXiv:quant-ph/03110091, doi:10.1016/j.physleta.2004.02.042.
- [17] G. J. Pryde, J. L. O’Brien, A. G. White, T. C. Ralph, and H. M. Wiseman. Measurement of quantum weak values of photon polarization. *Phys. Rev. Lett.*, 94:220405, 2005. arXiv:quant-ph/0412204, doi:10.1103/PhysRevLett.94.220405.
- [18] O. Hosten and P. Kwiat. Observation of the Spin Hall Effect of Light via Weak Measurements. *Science*, 319(5864):787, 2008. doi:10.1126/science.1152697.
- [19] 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. arXiv:0906.4828, doi:10.1103/PhysRevLett.102.173601.
- [20] 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. arXiv:0910.2410, doi:10.1103/PhysRevA.80.041803.
- [21] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell. Precision frequency measurements with interferometric weak values. *Phys. Rev. A*, 82:063822, 2010. arXiv:1101.1464, doi:10.1103/PhysRevA.82.063822.
- [22] 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(8):1479–1481, 2011. arXiv:1103.4426, doi:10.1364/OL.36.001479.
- [23] P. Egan and J. A. Stone. Weak-value thermostat with 0.2 mK precision. *Opt. Lett.*, 37(23):4991–4993, 2012. doi:10.1364/OL.37.004991.
- [24] L. J. Salazar-Serrano, D. Barrera, W. Amaya, S. Sales, V. Pruneri, J. Capmany, and J. P. Torres. Enhancement of the sensitivity of a temperature sensor based on fiber Bragg gratings via weak value amplification. *Opt. Lett.*, 40(17):3962–3965, 2015. doi:10.1364/OL.40.003962.
- [25] X. Xu, Y. Kedem, K. Sun, L. Vaidman, C. Li, and G. Guo. Phase Estimation with Weak Measurement Using a White Light Source. *Phys. Rev. Lett.*, 111:033604, 2013. arXiv:1306.4768, doi:10.1103/PhysRevLett.111.033604.
- [26] G. I. Viza, J. Martínez-Rincón, G. A. Howland, H. Frostig, I. Shomroni, B. Dayan, and J. C. Howell. Weak-values technique for velocity measurements. *Opt. Lett.*, 38(16):2949–2952, 2013. arXiv:1304.0029, doi:10.1364/OL.38.002949.

- [27] C. Wu, J. Zhang, Y. Xie, B. Ou, T. Chen, W. Wu, and P. Chen. Scheme and experimental demonstration of fully atomic weak-value amplification. *Phys. Rev. A*, 100:062111, 2019. arXiv:1811.06170, doi:10.1103/PhysRevA.100.062111.
- [28] L. Xu, L. Luo, H. Wu, Z. Luo, Z. Zhang, H. Shi, T. Chang, P. Wu, C. Du, and Hong-Liang Cui. Measurement of Chiral Molecular Parameters Based on a Combination of Surface Plasmon Resonance and Weak Value Amplification. *ACS Sensors*, 5(8):2398, 2020. doi:10.1021/acssensors.0c00346.
- [29] T. Koike and S. Tanaka. Limits on amplification by Aharonov-Albert-Vaidman weak measurement. *Phys. Rev. A*, 84:062106, 2011. arXiv:1108.2050, doi:10.1103/PhysRevA.84.062106.
- [30] S. Tanaka and N. Yamamoto. Information amplification via postselection: A parameter-estimation perspective. *Phys. Rev. A*, 88:042116, 2013. arXiv:1306.2409, doi:10.1103/PhysRevA.88.042116.
- [31] G. C. Knee and E. M. Gauger. When amplification with weak values fails to suppress technical noise. *Phys. Rev. X*, 4:011032, 2014. arXiv:1306.6321, doi:10.1103/PhysRevX.4.011032.
- [32] A. Nishizawa, K. Nakamura, and M. Fujimoto. Weak-value amplification in a shot-noise-limited interferometer. *Phys. Rev. A*, 85:062108, 2012. arXiv:1201.6039, doi:10.1103/PhysRevA.85.062108.
- [33] Y. Kedem. Using technical noise to increase the signal-to-noise ratio of measurements via imaginary weak values. *Phys. Rev. A*, 85:060102(R), 2012. arXiv:1111.5523, doi:10.1103/PhysRevA.85.060102.
- [34] 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. arXiv:1309.5011, doi:10.1103/PhysRevX.4.011031.
- [35] D. R. M. Arvidsson-Shukur, N. Y. Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and Seth Lloyd. Quantum advantage in postselected metrology. *Nat. Commun.*, 11(1):3775, 2020. arXiv:1903.02563, doi:10.1038/s41467-020-17559-w.
- [36] J. Lee and I. Tsutsui. Merit of amplification by weak measurement in view of measurement uncertainty. *Quantum Stud.: Math. Found.*, 1(1):65, 2014. arXiv:1305.2721, doi:10.1007/s40509-014-0002-x.
- [37] S. Pang, J. R. G. Alonso, T. A. Brun, and A. N. Jordan. Protecting weak measurements against systematic errors. *Phys. Rev. A*, 94:012329, 2016. arXiv:1605.09040, doi:10.1103/PhysRevA.94.012329.
- [38] 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. arXiv:1305.4520, doi:10.1103/PhysRevA.88.023821.
- [39] S. Pang, J. Dressel, and T. A. Brun. Entanglement-Assisted Weak Value Amplification. *Phys. Rev. Lett.*, 113:030401, 2014. arXiv:1401.5887, doi:10.1103/PhysRevLett.113.030401.

- [40] C. Fang, J. Huang, Y. Yu, Q. Li, and G. Zeng. Ultra-small time-delay estimation via a weak measurement technique with post-selection. *J. Phys. B. At. Mol. Opt. Phys.*, 49(17):175501, 2016. arXiv:1509.04003, doi:10.1088/0953-4075/49/17/175501.
- [41] L. J. Salazar-Serrano, D. Janner, N. Brunner, V. Pruneri, and J. P. Torres. Measurement of sub-pulse-width temporal delays via spectral interference induced by weak value amplification. *Phys. Rev. A*, 89:012126, 2014. arXiv:1312.4469, doi:10.1103/PhysRevA.89.012126.
- [42] L. Luo, L. Xie, J. Qiu, Xinxing Zhou, X. Liu, Z. Li, Y. He, Z. Zhang, and H. Sun. Simultaneously precise estimations of phase and amplitude variations based on weak-value amplification. *Appl. Phys. Lett.*, 114(11):111104, 2019. doi:10.1063/1.5083995.
- [43] K. Ogawa, H. Kobayashi, and A. Tomita. Operational formulation of weak values without probe systems. *Phys. Rev. A*, 101:042117, 2020. arXiv:1912.10222, doi:10.1103/PhysRevA.101.042117.
- [44] A. Nishizawa. Weak-value amplification beyond the standard quantum limit in position measurements. *Phys. Rev. A*, 92:032123, 2015. arXiv:1504.04777, doi:10.1103/PhysRevA.92.032123.
- [45] S. Huang and G. S. Agarwal. Weak value amplification of atomic cat states. *New J. Phys.*, 17(9):093032, 2015. arXiv:1501.02359, doi:10.1088/1367-2630/17/9/093032.
- [46] K. Kawana and D. Ueda. Amplification of gravitational motion via Quantum weak measurement. *Prog. Theor. Exp. Phys.*, 2019(4):041A01, 2019. arXiv:1804.09505, doi:10.1093/ptep/ptz019.
- [47] D. Ueda and T. Kitahara. Novel approach to neutron electric dipole moment search using weak measurement. 2020. arXiv:2002.11731, doi:10.1088/1361-6455/abc5a0.
- [48] K. Nakamura and M. Fujimoto. Extension of the input-output relation for a michelson interferometer to arbitrary coherent-state light sources: Gravitational-wave detector and weak-value amplification. *Ann. Phys.*, 392:71, 2018. arXiv:1705.09768, doi:10.1016/j.aop.2018.03.009.
- [49] B. Xia, J. Huang, C. Fang, H. Li, and G. Zeng. High-precision multiparameter weak measurement with hermite-gaussian pointer. *Phys. Rev. Appl.*, 13:034023, 2020. doi:10.1103/PhysRevApplied.13.034023.
- [50] T. Denkmayr, H. Geppert, S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa. Observation of a quantum Cheshire Cat in a matter-wave interferometer experiment. *Nat. Commun.*, 5(1):4492, 2014. arXiv:1312.3775, doi:10.1038/ncomms5492.
- [51] T. Denkmayr, J. Dressel, H. Geppert-Kleinrath, Y. Hasegawa, and S. Sponar. Weak values from strong interactions in neutron interferometry. *Physica B*, 551:339, 2018. The 11th International Conference on Neutron Scattering (ICNS 2017). doi:10.1016/j.physb.2018.04.014.

- [52] M. Zhang. Weak values obtained from mass-energy equivalence. *Phys. Rev. A*, 95:012114, 2017. arXiv:1601.06077, doi:10.1103/PhysRevA.95.012114.
- [53] A. Go. Observation of bell inequality violation in b mesons. *J. Mod. Opt.*, 51(6-7):991–998, 2004. arXiv:quant-ph/0310192, doi:10.1080/09500340408233614.
- [54] Y. Mori, J. Lee, and I. Tsutsui. On the validity of weak measurement applied for precision measurement. *J. Phys. B. At. Mol. Opt. Phys.*, 53(1):015501, 2019. arXiv:1901.06831, doi:10.1088/1361-6455/ab5200.
- [55] S. Higashino, Y. Mori, Y. Takubo, A. Ishikawa, T. Higuchi, and I. Tsutsui. Novel Approach to Measuring CP Violation in B Meson Decays: A Case Study of Weak Measurement in High Energy Physics. 2020. arXiv:2011.07560.
- [56] S. G. Rabinovich. *Measurement Errors and Uncertainties Theory and Practice*. Springer, third edition, 2005.
- [57] P. A. Zyla et al. Review of Particle Physics. *Prog. Theor. Exp. Phys.*, 2020(8):083C01, 2020. doi:10.1093/ptep/ptaa104.
- [58] J. Neyman. On the problem of confidence intervals. *Ann. Math. Stat.*, 6(3):111–116, 1935. doi:10.1214/aoms/1177732585.
- [59] Y. Aharonov and A. Botero. Quantum averages of weak values. *Phys. Rev. A*, 72:052111, 2005. arXiv:quant-ph/0503225, doi:10.1103/PhysRevA.72.052111.
- [60] M. Ban. Conditional average in a quantum system with postselection. *Quantum Stud.: Math. Found.*, 2(3):263, 2015. doi:10.1007/s40509-015-0043-9.
- [61] I. Shomroni, O. Bechler, S. Rosenblum, and B. Dayan. Demonstration of Weak Measurement Based on Atomic Spontaneous Emission. *Phys. Rev. Lett.*, 111:023604, 2013. arXiv:1304.2912, doi:10.1103/PhysRevLett.111.023604.
- [62] M. Onoda, S. Murakami, and N. Nagaosa. Hall effect of light. *Phys. Rev. Lett.*, 93:083901, 2004. arXiv:cond-mat/0405129, doi:10.1103/PhysRevLett.93.083901.
- [63] S. Chen and C. Mi and W. Wu and W. Zhang and W. Shu and H. Luo and S. Wen. Weak-value amplification for weyl-point separation in momentum space. *New J. Phys.*, 20(10):103050, 2018. arXiv:1802.06884, doi:10.1088/1367-2630/aae2d5.
- [64] S. Carrasco and M. Orszag. Weak-value amplification of photon-number operators in the optomechanical interaction. *Phys. Rev. A*, 99:013801, 2019. arXiv:1804.10252, doi:10.1103/PhysRevA.99.013801.
- [65] 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. arXiv:1504.08210, doi:10.1103/PhysRevLett.114.170801.
- [66] J. Harris, R. W. Boyd, and J. S. Lundeen. Weak value amplification can outperform conventional measurement in the presence of detector saturation. *Phys. Rev. Lett.*, 118:070802, 2017. arXiv:1612.04327, doi:10.1103/PhysRevLett.118.070802.
- [67] D. F. Walls and G. J. Milburn. *Quantum Optics*. Springer, second edition, 2008.

- [68] R. Kumar, E. Barrios, C. Kupchak, and A. I. Lvovsky. Experimental Characterization of Bosonic Creation and Annihilation Operators. *Phys. Rev. Lett.*, 110:130403, 2013. [arXiv:1210.1150](#), [doi:10.1103/PhysRevLett.110.130403](#).
- [69] J. D. Jackson. *Classical Electrodynamics*. Wiley, third edition, 1998.
- [70] J. C. Howell, D. J. Starling, P. B. Dixon, P. K. Vudyasetu, and A. N. Jordan. Interferometric weak value deflections: Quantum and classical treatments. *Phys. Rev. A*, 81:033813, 2010. [arXiv:0906.4832](#), [doi:10.1103/PhysRevA.81.033813](#).
- [71] J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, second edition, 1996.
- [72] N. Morisawa. 弱値増幅による光のスピンホール効果の測定不確かさについて (On the measurement uncertainty via weak-value amplification for the Spin Hall Effect of Light). Master's thesis, The University of Tokyo, 2016.
- [73] J. Sinclair, M. Hallaji, A. M. Steinberg, J. Tollaksen, and A. N. Jordan. Weak-value amplification and optimal parameter estimation in the presence of correlated noise. *Phys. Rev. A*, 96:052128, 2017. [arXiv:1703.01496](#), [doi:10.1103/PhysRevA.96.052128](#).
- [74] T. Likhomanenko, D. Derkach, and A. Rogozhnikov. Inclusive Flavour Tagging Algorithm. *J. Phys. Conf. Ser.*, 762:012045, 2016. [arXiv:1705.08707](#), [doi:10.1088/1742-6596/762/1/012045](#).
- [75] F. Abudinén et al. First flavor tagging calibration using 2019 Belle II data. 2020. [arXiv:2008.02707](#).
- [76] B. Grinstein, Y. Grossman, Z. Ligeti, and D. Pirjol. Photon polarization in $B \rightarrow X\gamma$ in the standard model. *Phys. Rev. D*, 71:011504, 2005. [arXiv:hep-ph/0412019](#), [doi:10.1103/PhysRevD.71.011504](#).
- [77] H. Kakuno et al. Neutral B flavor tagging for the measurement of mixing-induced CP violation at Belle. *Nucl. Instrum. Meth. A*, 533. [arXiv:hep-ex/0403022](#), [doi:10.1016/j.nima.2004.06.159](#).
- [78] T. Horiguchi et al. Evidence for Isospin Violation and Measurement of CP Asymmetries in $B \rightarrow K^*(892)\gamma$. *Phys. Rev. Lett.*, 119:191802, 2017. [arXiv:1707.00394](#), [doi:10.1103/PhysRevLett.119.191802](#).
- [79] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71(2):1554, 2011. [Erratum: *Eur. Phys. J. C* **73**, 2501 (2013)]. [arXiv:1007.1727](#), [doi:10.1140/epjc/s10052-011-1554-0](#).
- [80] Y. Ushiroda et al. Time-dependent CP asymmetries in $B^0 \rightarrow K_S^0\pi^0\gamma$ transitions. *Phys. Rev. D*, 74:111104, 2006. [arXiv:hep-ex/0608017](#), [doi:10.1103/PhysRevD.74.111104](#).
- [81] Y. S. Amhis et al. Averages of b -hadron, c -hadron, and τ -lepton properties as of 2018. Technical report, 2019. [arXiv:1909.12524](#).
- [82] R. A. Bertlmann and B. C. Hiesmayr. Strangeness measurements of kaon pairs, CP violation and Bell inequalities. [arXiv:hep-ph/0609251](#).

- [83] G. D'ambrosio and G. Isidori. CP VIOLATION IN KAON DECAYS. *Int. J. Mod. Phys. A*, 13(01):1, 1998. [arXiv:hep-ph/9611284](#), [doi:10.1142/S0217751X98000020](#).
- [84] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, first edition, 1993.
- [85] K. Igi and H. Kawai. 量子力学 *I, II*. Kohdansha, first edition, 1994.
- [86] P. D. Mannheim. CPT Symmetry Without Hermiticity. *PoS, ICHEP2016:347*, 2016. [arXiv:1611.02100](#), [doi:10.22323/1.282.0347](#).