Complex counterpart of variance in quantum measurements for pre- and postselected systems

Kazuhisa Ogawa,[1,*] Natsuki Abe,[1] Hirokazu Kobayashi,[2] and Akihisa Tomita[1]
[1]*Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan*
[2]*School of System Engineering, Kochi University of Technology, Tosayamada-cho, Kochi 782-8502, Japan*

(Received 14 March 2021; accepted 28 June 2021; published 21 July 2021)

The variance of an observable in a preselected quantum system, which is always real and non-negative, appears as an increase in the probe wave packet width in indirect measurements. Extending this framework to pre- and postselected systems, we formulate a complex-valued counterpart of the variance called "weak variance." In our formulation, the real and imaginary parts of the weak variance appear as changes in the probe wave packet width in the vertical-horizontal and diagonal-antidiagonal directions, respectively, on the quadrature phase plane. Using an optical system, we experimentally demonstrate these changes in the probe wave packet width caused by the real negative and purely imaginary weak variances. Furthermore, we show that the weak variance can be expressed as the variance of the weak-valued probability distribution in pre- and postselected systems. These operational and statistical interpretations support the rationality of formulating the weak variance as a complex counterpart of the variance in pre- and postselected systems.

DOI: 10.1103/PhysRevResearch.3.033077

I. INTRODUCTION

The outcomes of quantum measurements show probabilistic behavior. This characteristic, which is not observed in classical systems, has been the root of many fundamental arguments in quantum theory [1]. In the quantum measurement of an observable $\hat{A}$, the probabilistic behavior of its measurement outcomes is characterized by measurement statistics such as expectation value $\langle\hat{A}\rangle$ and variance $\sigma^2(\hat{A})$. These values are generally measured using an indirect measurement method [2]. In indirect measurements, the target system to be measured is coupled with an external probe system through von Neumann interaction. Regardless of the coupling strength, the expectation value $\langle\hat{A}\rangle$ and variance $\sigma^2(\hat{A})$ in the target system are obtained from the displacement of the probe wave packet and the increase in its width, respectively. In other words, the probe wave packet in an indirect measurement serves as the interface that displays the probabilistic characteristics of the target system.

Interestingly, when the target system is further postselected, the displacement of the probe wave packet differs from $\langle\hat{A}\rangle$. In particular, when the coupling strength is weak (weak measurement setup), the probe displacement is given by $\mathrm{Re}\langle\hat{A}\rangle_{\mathrm{w}}$, where $\langle\hat{A}\rangle_{\mathrm{w}} := \langle \mathrm{f}|\hat{A}|i\rangle/\langle \mathrm{f}|i\rangle$ in the pre- and postselected system $\{|i\rangle, |\mathrm{f}\rangle\}$ is called weak value [3]. $\langle\hat{A}\rangle_{\mathrm{w}}$ is complex in general and can exceed the spectral range of $\hat{A}$. By regarding the weak value as a complex counterpart of the expectation value in the pre- and postselected system, new approaches to fundamental problems in quantum mechanics involving pre- and postselection have been investigated, such as various quantum paradoxes [4–10], understanding of the violation of Bell's inequality using negative probabilities [11], the relationship between disturbance and complementarity in quantum measurements [12–14], verification of the uncertainty relations [15–17], observation of Bohmian trajectories [18,19], and demonstration of the violation of macrorealism [20,21].

Similar to the relation between the weak value and expectation value, does there exist a counterpart of the variance in pre- and postselected systems? To answer this question, we consider the function of the probe wave packet in indirect measurement as an interface that displays the characteristics of the target system. As mentioned earlier, the variance $\sigma^2(\hat{A})$ in a preselected system manifests as an increase in the probe wave packet width in indirect measurement, and owing to the non-negativity of the variance, the wave packet width never decreases. However, for pre- and postselected systems, any counterparts of the variance cannot be observed in the typical framework of the weak measurement [3], in which the probe wave packet width does not change because the second- and higher order terms of the coupling strength are ignored. Here, we focus on the recent studies reporting that when considering the second- and higher order terms of the coupling strength, the probe wave packet width can not only increase but also decrease under appropriate pre- and postselection conditions [22,23]. If these reported phenomena are interpreted to result from a counterpart of the variance in pre- and postselected systems, it may be possible to formulate an effective variancelike quantity that can be negative.

In this study, we investigate the general changes in the width of the probe wave packet during indirect measurements of pre- and postselected systems. We then formulate a

*ogawak@ist.hokudai.ac.jp

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

counterpart of the variance in these systems. This counterpart, denoted here as *weak variance*, can indeed be negative and manifests as the decrease in the probe wave packet width. Moreover, the weak variance is generally complex and can be understood by observing the changes in the probe wave packet width on the quadrature phase plane. To demonstrate this phenomenon, we conducted an optical experiment for observing the changes in the beam packet width in proportion to the real and imaginary parts of the complex weak variance. In addition, to clarify the concept of weak variance, we express the weak variance as the second-order moment of the weak-valued probability distribution [4–13,24–26], which is a quasiprobability distribution in pre- and postselected systems. Based on the agreement between the operational and statistical interpretations, our weak variance can be considered a more reasonable definition of a complex counterpart of the variance in pre- and postselected systems than previous formulations [27–39]. Furthermore, we formulate a counterpart of the higher order moment and investigate its operational and statistical meanings and applications.

II. WEAK VARIANCE APPEARING IN INDIRECT MEASUREMENT FOR PRE- AND POSTSELECTED SYSTEMS

The indirect measurements have been made with a Gaussian probe as shown in Figs. 1(a) and 1(d). After reviewing these measurements, we explain the appearance of a complex weak variance in the pre- and postselected system. The target system to be measured and the probe system are preselected in states $|i\rangle$ and $|\phi\rangle$, respectively. The initial probe state $|\phi\rangle$ can be expanded as $|\phi\rangle = \int_{-\infty}^{\infty} dX \phi(X)|X\rangle$, where the wave function $\phi(X)$ is the Gaussian distribution $\phi(X) = \pi^{-1/4}\exp(-X^2/2)$ and X is a dimensionless variable.[1] The observable of the dimensionless position $\hat{X}$ can be spectrally decomposed as $\hat{X} = \int_{-\infty}^{\infty} dX X|X\rangle\langle X|$. The time evolution by the interaction Hamiltonian $\hat{A}\otimes\hat{K}$ is represented by the unitary operator $\hat{U}(\theta) = \exp(-i\theta\hat{A}\otimes\hat{K})$, where $\hat{A} = \sum_j a_j \hat{\Pi}_j$ is the observable to be measured in the target system, a_j is an eigenvalue of $\hat{A}$, $\hat{\Pi}_j$ is the projector onto the eigenspace of $\hat{A}$ belonging to eigenvalue a_j, $\hat{K}$ is the canonical conjugate observable of $\hat{X}$ satisfying $[\hat{X},\hat{K}] = i\hat{1}$, and θ is a parameter with the reciprocal dimension of $\hat{A}$. The coupling strength is characterized by $\theta\|\hat{A}\|$, where $\|\hat{A}\|$ is the largest eigenvalue of $\hat{A}$: if $\theta\|\hat{A}\| \gg 1$ ($\ll 1$), the coupling is considered strong (weak).

Let us consider an indirect measurement of the observable $\hat{A}$, as shown in Fig. 1(a). Suppose that measurement $\hat{X}$ in the probe system is made to the state after the interaction, $|\Omega\rangle := \exp(-i\theta\hat{A}\otimes\hat{K})|i\rangle|\phi\rangle$. The probability distribution $P(X)$ of obtaining the result X is

$$P(X) = |\langle X|\Omega\rangle|^2 = \sum_j p_j |\phi(X-\theta a_j)|^2, \tag{1}$$

[1]The dimensionless variable X is obtained by dividing the position variable x by the standard deviation σ of the wave function $\pi^{-1/4}\sigma^{-1/2}\exp[-x^2/(2\sigma^2)]$.

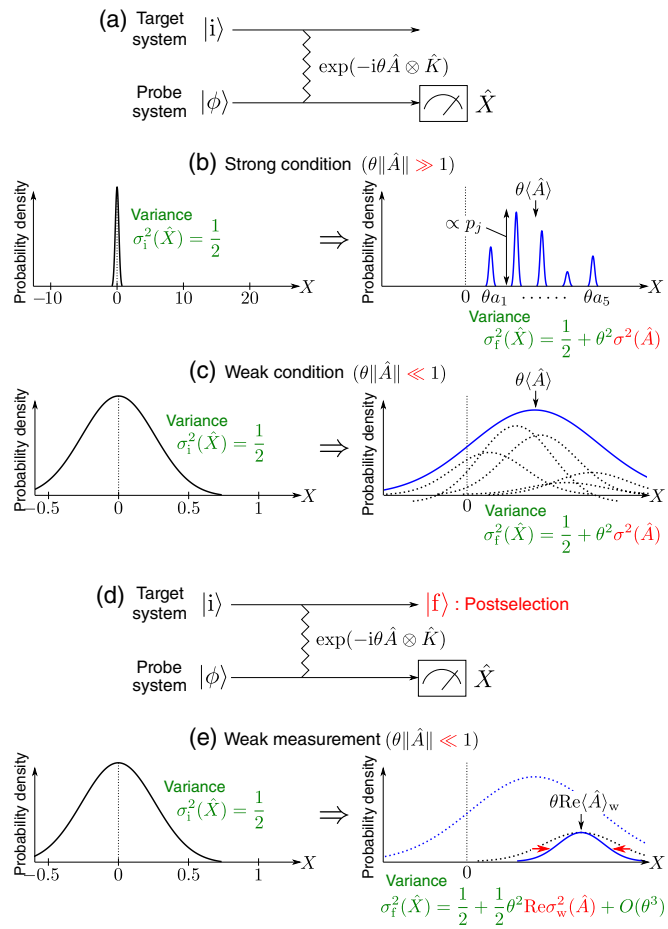

FIG. 1. (a) Quantum circuit of indirect measurements of the preselected system $|i\rangle$. (b) Change in the probe wave packet caused by interactions in the quantum circuit (a) under the strong coupling condition $(\theta\|\hat{A}\| \gg 1)$. The distribution of probe wave packets after the interaction reproduces the probability distribution of the outcomes of projective measurements of $\hat{A}$ in $|i\rangle$. (c) Change in the probe wave packet under the weak coupling condition $(\theta\|\hat{A}\| \ll 1)$, where the horizontal axis has been rescaled from that of (b). The variance of the probe wave packet after the interaction increases in proportion to the variance $\sigma^2(\hat{A})$ and never decreases. (d) Quantum circuit of indirect measurements of the pre- and postselected system $\{|i\rangle, |f\rangle\}$ (weak measurement setup). (e) Change in the probe wave packet in the weak measurement circuit (d). The real part of the weak variance appears in the variance change of the probe wave packet after the postselection. Unlike the preselected system (c), the pre- and postselected system admits a narrowed variance of the probe wave packet when the real part of the weak variance becomes negative.

where $p_j := \langle i|\hat{\Pi}_j|i\rangle$ is the projection probability of $|i\rangle$ onto $\hat{\Pi}_j$. If the coupling is strong $(\theta\|\hat{A}\| \gg 1)$, the wave packet $|\phi(X-\theta a_j)|^2$ for each j is well separated from other wave packets, and $P(X)$ reproduces the probability distribution $\{p_j\}_j$ [Fig. 1(b)]. However, if the coupling is weak $(\theta\|\hat{A}\| \ll 1)$, the wave packets overlap and $P(X)$ does not reproduce $\{p_j\}_j$ [Fig. 1(c)]. Nevertheless, regardless of the coupling strength, the statistics of $\hat{A}$ in the target system $|i\rangle$, such as the expectation value $\langle\hat{A}\rangle$ and the variance $\sigma^2(\hat{A})$, can be acquired from the changes in the probe distribution $P(X)$. The

expectation value and variance of X in $P(X)$ are respectively expressed as

$$\langle\hat{X}\rangle_{\rm f} = \langle\hat{X}\rangle_i + \theta\langle\hat{A}\rangle, \quad \sigma_{\rm f}^2(\hat{X}) = \sigma_i^2(\hat{X}) + \theta^2\sigma^2(\hat{A}), \tag{2}$$

where $\langle\hat{X}\rangle_i$ and $\sigma_i^2(\hat{X})$ are the expectation value and variance of $\hat{X}$ in the initial probe state $|\phi\rangle$, respectively. In this case, $\langle\hat{X}\rangle_i = 0$ and $\sigma_i^2(\hat{X}) = 1/2$. Therefore the expectation value $\langle\hat{A}\rangle$ and variance $\sigma^2(\hat{A})$ can be measured under both strong and weak coupling conditions. Here, we stress that after the interaction, the variance of the probe wave packet $\sigma_{\rm f}^2(\hat{X})$ never decreases because the variance $\sigma^2(\hat{A})$ is non-negative.

We next consider that the target system is pre- and postselected in states $|i\rangle$ and $|{\rm f}\rangle$, respectively [Fig. 1(d)]. The non-normalized state of the probe system after the postselection $|\tilde{\phi}_{\rm f}\rangle := \langle{\rm f}|\Omega\rangle$ is represented as

$$|\tilde{\phi}_{\rm f}\rangle = \langle{\rm f}|i\rangle\left(\hat{1} - i\theta\langle\hat{A}\rangle_{\rm w}\hat{K} - \frac{\theta^2}{2}\langle\hat{A}^2\rangle_{\rm w}\hat{K}^2\right)|\phi\rangle + O(\theta^3). \tag{3}$$

The expectation value of $\hat{X}$ in the non-normalized state $|\tilde{\phi}_{\rm f}\rangle$ is $\langle\hat{X}\rangle_{\rm f} = \langle\tilde{\phi}_{\rm f}|\hat{X}|\tilde{\phi}_{\rm f}\rangle/\langle\tilde{\phi}_{\rm f}|\tilde{\phi}_{\rm f}\rangle = {\rm Re}\langle\hat{A}\rangle_{\rm w}\theta + O(\theta^3)$. The real part of the weak value $\langle\hat{A}\rangle_{\rm w} = \langle{\rm f}|\hat{A}|i\rangle/\langle{\rm f}|i\rangle$ appears in the displacement of the probe wave packet, as previously reported for weak measurements [3]. The imaginary part of the weak value is observed in the displacement of the probe wave packet in the $\hat{K}$ basis: $\langle\hat{K}\rangle_{\rm f} = {\rm Im}\langle\hat{A}\rangle_{\rm w}\theta + O(\theta^3)$ [40]. By introducing the generalized position operator $\hat{M} := \hat{X}\cos\alpha + \hat{K}\sin\alpha$ ($\alpha \in [0, 2\pi)$), these relations can be summarized as

$$\langle\hat{M}\rangle_{\rm f} = (\cos\alpha\,{\rm Re}\langle\hat{A}\rangle_{\rm w} + \sin\alpha\,{\rm Im}\langle\hat{A}\rangle_{\rm w})\theta + O(\theta^3). \tag{4}$$

The real and imaginary parts of the weak value are also obtained by setting the interaction Hamiltonian to $\hat{A}\otimes\hat{M}$ and measuring the expectation values of $\hat{X}$ and $\hat{K}$ for the postselected probe state [41].

Now let us examine the change in the probe wave packet width. The variance of $\hat{X}$ for $|\tilde{\phi}_{\rm f}\rangle$ is calculated as

$$\begin{aligned}\sigma_{\rm f}^2(\hat{X}) &= \langle\hat{X}^2\rangle_{\rm f} - \langle\hat{X}\rangle_{\rm f}^2 \\ &= \sigma_i^2(\hat{X}) + \tfrac{1}{2}{\rm Re}\big(\langle\hat{A}^2\rangle_{\rm w} - \langle\hat{A}\rangle_{\rm w}^2\big)\theta^2 + O(\theta^3).\end{aligned} \tag{5}$$

The real part of the variancelike quantity appears in the quadratic term of θ, which is ignored in the conventional weak measurement context. We define this quantity as the *weak variance* $\sigma_{\rm w}^2(\hat{A})$ of $\hat{A}$:

$$\sigma_{\rm w}^2(\hat{A}) := \langle\hat{A}^2\rangle_{\rm w} - \langle\hat{A}\rangle_{\rm w}^2. \tag{6}$$

The real part of the weak variance is similar to normal variance in that it appears as a change in the probe wave packet width in the $\hat{X}$ basis [Eq. (2)]. However, unlike the normal variance, the weak variance can be negative, in which case the wave packet width then decreases as shown in Fig. 1(e). The decrease in the probe wave packet width reported in previous studies [22,23] can be reinterpreted as the effect of the negative weak variance.

We next consider the appearance of the imaginary part of the weak variance. The variance of the generalized position operator $\hat{M}$ for $|\tilde{\phi}_{\rm f}\rangle$ is calculated as (see Appendix A for a detailed analysis of mixed pre- and postselected states of the target system)

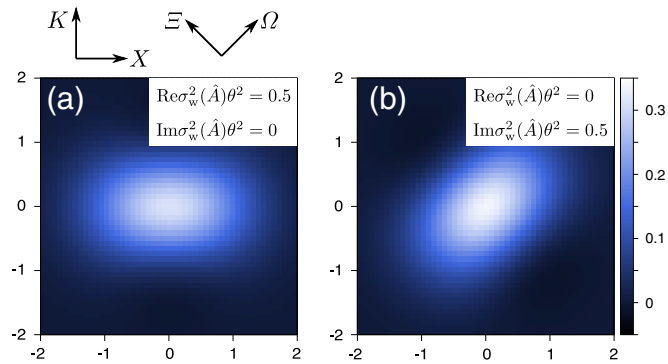

FIG. 2. Wigner functions of the normalized probe states after postselection, $|\tilde{\phi}_{\rm f}\rangle/\||\tilde{\phi}_{\rm f}\rangle\|$, assuming $\langle\hat{A}\rangle_{\rm w} = 0$ and neglecting $O(\theta^3)$ in Eq. (3). The horizontal and vertical axes represent the observables $\hat{X}$ and $\hat{K}$, respectively, and the 45° and 135° axes represent the observables $\hat{\Omega}$ and $\hat{\Xi}$, respectively. (a) When ${\rm Re}\sigma_{\rm w}^2(\hat{A})\theta^2 = 0.5$ and ${\rm Im}\sigma_{\rm w}^2(\hat{A})\theta^2 = 0$, the wave packet spreads along the X axis and narrows along the K axis. (b) When ${\rm Re}\sigma_{\rm w}^2(\hat{A})\theta^2 = 0$ and ${\rm Im}\sigma_{\rm w}^2(\hat{A})\theta^2 = 0.5$, the wave packet spreads along the Ω axis and narrows along the Ξ axis.

$$\begin{aligned}\sigma_{\rm f}^2(\hat{M}) = \tfrac{1}{2} + \tfrac{1}{2}[\cos(2\alpha){\rm Re}\sigma_{\rm w}^2(\hat{A}) + \sin(2\alpha){\rm Im}\sigma_{\rm w}^2(\hat{A})]\theta^2 \\ + O(\theta^3).\end{aligned} \tag{7}$$

This equation indicates that the real and imaginary parts of the weak variance appear in the changes in the probe wave packet width in different measurement bases. For example, when choosing $\hat{M} = \hat{K}$ ($\alpha = \pi/2$), the variance of $\hat{K}$ in state $|\tilde{\phi}_{\rm f}\rangle$ is given as $\sigma_{\rm f}^2(\hat{K}) = \sigma_i^2(\hat{K}) - (1/2){\rm Re}\sigma_{\rm w}^2(\hat{A})\theta^2 + O(\theta^3)$; that is, ${\rm Re}\sigma_{\rm w}^2(\hat{A})$ also appears in the width change of the wave packet in the $\hat{K}$ basis. To clarify these relations, we plot them on the quadrature phase plane [Fig. 2(a)]. When ${\rm Re}\sigma_{\rm w}^2(\hat{A}) > 0$, the wave packet spreads along the X (horizontal) axis, while it narrows along the K (vertical) axis. This relationship satisfies the Kennard-Robertson uncertainty relation [42,43] up to the quadratic of θ: $\sigma_{\rm f}^2(\hat{X})\sigma_{\rm f}^2(\hat{K}) = 1/4 + O(\theta^3)$.

However, when $\alpha = \pi/4$, the measured observable becomes $\hat{M} = (\hat{X} + \hat{K})/\sqrt{2} =: \hat{\Omega}$, which corresponds to the 45° axis in the quadrature phase plane of Fig. 2. Through the relation $\sigma_{\rm f}^2(\hat{\Omega}) = \sigma_i^2(\hat{\Omega}) + (1/2){\rm Im}\sigma_{\rm w}^2(\hat{A})\theta^2 + O(\theta^3)$, the imaginary part of the weak variance ${\rm Im}\sigma_{\rm w}^2(\hat{A})$ can be observed as the change in the probe wave packet width in the $\hat{\Omega}$ basis. Moreover, when $\alpha = 3\pi/4$, the measured observable becomes $\hat{M} = (-\hat{X} + \hat{K})/\sqrt{2} =: \hat{\Xi}$, which is the canonical conjugate of $\hat{\Omega}$. This observable satisfies $[\hat{\Omega}, \hat{\Xi}] = i\hat{1}$ and corresponds to the 135° axis in the quadrature phase plane of Fig. 2. Through the relation $\sigma_{\rm f}^2(\hat{\Xi}) = \sigma_i^2(\hat{\Xi}) - (1/2){\rm Im}\sigma_{\rm w}^2(\hat{A})\theta^2 + O(\theta^3)$, ${\rm Im}\sigma_{\rm w}^2(\hat{A})$ also appears in the width change of the wave packet in the $\hat{\Xi}$ basis. These relations are represented on the quadrature phase plane in Fig. 2(b). When ${\rm Im}\sigma_{\rm w}^2(\hat{A}) > 0$, the wave packet spreads along the Ω (45°) axis, while it narrows along the Ξ (135°) axis. This relationship also satisfies the Kennard-Robertson uncertainty relation up to the quadratic of θ: $\sigma_{\rm f}^2(\hat{\Omega})\sigma_{\rm f}^2(\hat{\Xi}) = 1/4 + O(\theta^3)$.

We note that our formalism of the weak variance is reconciled with the claim by Vaidman *et al.* [44] that the weak value

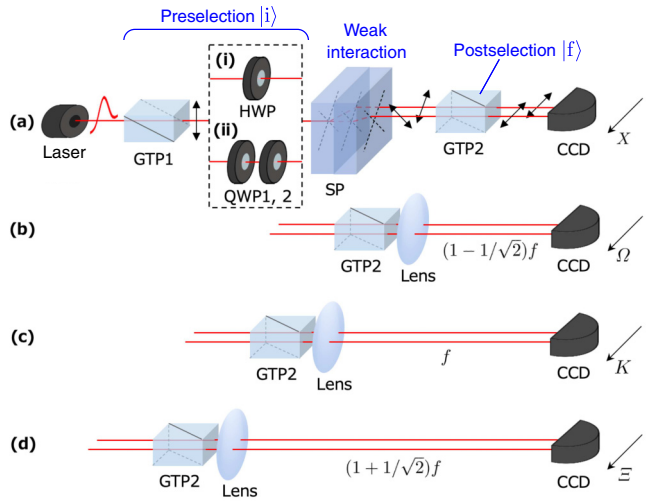

FIG. 3. Experimental setup for observing weak variances. GTP: Glan-Thompson prism; HWP: half-wave plate; QWP: quarter-wave plate; SP: Savart plate; CCD: charge-coupled device. (a) Experimental setup for measuring the probe system in the $\hat{X}$ basis. In the preselection, the HWP and two QWPs are used to prepare the weak variance to be (i) negative real and (ii) positive purely imaginary, respectively. [(b), (c), and (d)] Experimental setups for measuring the probe system in the $\hat{\Omega}$, $\hat{K}$, and $\hat{\Xi}$ bases, respectively. The lens (focal length $f = 1$ m) and free-space propagation perform a fractional Fourier transform on the transverse distribution of the beam.

is a robust property of the system and, unlike an expectation value, is not a statistical property. This can be understood by considering the changes in the probe wave packet on the quadrature phase plane. Considering up to the first order of θ, the probe wave packet after the postselection is displaced by $(\theta \mathrm{Re}\langle\hat{A}\rangle_{\mathrm{w}}, \theta \mathrm{Im}\langle\hat{A}\rangle_{\mathrm{w}})$, but its shape is unchanged. The probe wave packet remains a pure state and can be specified with one complex parameter, the weak value. In this sense, the weak value is not statistical and is a robust property of the system like the eigenvalue. Considering up to the quadrature of θ, the probe wave packet is squeezed/antisqueezed in the direction specified by the weak variance, as explained above. The probe wave packet remains a pure state while its shape changes; therefore, the weak variance does not represent the statistical nature of the probabilistic mixture of probe wave packets and hence is consistent with the claim of Vaidman *et al.* Considering higher orders of θ, the probe wave packet undergoes a non-Gaussian transformation specified by the higher order weak moments $\langle\hat{A}^n\rangle_{\mathrm{w}}$, as described at the end of this paper.

III. EXPERIMENTAL DEMONSTRATION OF WEAK VARIANCES

To verify the effects of the weak variance, we experimentally observed the weak variance in the optical system shown in Fig. 3. In this setup, the target and probe systems were the polarization and transverse spatial modes, respectively, of the laser beam with a central wavelength of 780 nm (Menlo Systems C-fiber 780). The polarization mode was a two-state system spanned by (for example) the horizontal-vertical polarization basis $\{|\mathrm{H}\rangle, |\mathrm{V}\rangle\}$ or the diagonal (45°)-antidiagonal (135°) polarization basis $\{|\mathrm{D}\rangle := (|\mathrm{H}\rangle + |\mathrm{V}\rangle)/\sqrt{2}, |\mathrm{A}\rangle := (|\mathrm{H}\rangle - |\mathrm{V}\rangle)/\sqrt{2}\}$. The pre- and postselection $\{|i\rangle, |\mathrm{f}\rangle\}$ in the polarization mode was prepared using Glan-Thompson prisms (GTPs), a half-wave plate (HWP), and quarter-wave plates (QWPs). The initial transverse distribution of the beam's amplitude was prepared as a Gaussian distribution $\phi(X) = \pi^{-1/4}\exp(-X^2/2)$, where X is the dimensionless position variable normalized by the standard deviation of this distribution. The weak interaction $\exp(-i\theta\hat{A}\otimes\hat{K})$ was implemented using a Savart plate (SP), which comprises two orthogonal birefringent crystals (β-$\mathrm{BaB_2O_4}$, 1-mm thickness). In our setup, $\hat{A}$ was chosen as $\hat{A} = |\mathrm{D}\rangle\langle\mathrm{D}| - |\mathrm{A}\rangle\langle\mathrm{A}|$ and the SP transversely shifted the diagonally (antidiagonally) polarized beam by a distance of θ $(-\theta)$. The probe system was finally measured in the $\hat{X}$, $\hat{\Omega}$, $\hat{K}$, and $\hat{\Xi}$ bases. In the $\hat{X}$ basis, the transverse intensity distribution of the beam was measured using a charge-coupled device (CCD) camera (Teledyne Princeton Instruments ProEM-HS:512BX3), as shown in Fig. 3(a). The intensity measurements in the $\hat{\Omega}$, $\hat{K}$, and $\hat{\Xi}$ bases were implemented by fractional Fourier transforming (see Appendix B for details) the beam distribution using a lens (focal length $f = 1$ m) before $\hat{X}$ measurement using the CCD camera, as shown in Figs. 3(b)–3(d).

To independently verify the effects of the real and imaginary parts of the weak variance, we chose the pre- and postselected polarization states $\{|i\rangle, |\mathrm{f}\rangle\}$ giving (i) negative real and (ii) positive purely imaginary weak variances. The preselected state $|i\rangle$ in case (i) was prepared by rotating the fast axis of the HWP through angle ϑ_{H} from the vertical direction and passing the vertically polarized beam through the rotated HWP. The output state became $|i\rangle = \cos(2\vartheta_{\mathrm{H}} - \pi/4)|\mathrm{D}\rangle + \sin(2\vartheta_{\mathrm{H}} - \pi/4)|\mathrm{A}\rangle$. The postselected state was fixed at $|\mathrm{f}\rangle = |\mathrm{H}\rangle$. The weak value and weak variance respectively became the following real numbers:

$$\langle\hat{A}\rangle_{\mathrm{w}} = \frac{\cos(2\vartheta_{\mathrm{H}})}{\sin(2\vartheta_{\mathrm{H}})}, \quad \sigma_{\mathrm{w}}^2(\hat{A}) = -\frac{\cos(4\vartheta_{\mathrm{H}})}{\sin^2(2\vartheta_{\mathrm{H}})}. \tag{8}$$

The preselected state $|i\rangle$ in case (ii) was prepared by rotating the fast axis of QWP1 through angle ϑ_{Q} from the vertical direction and passing the vertically polarized beam through the rotated QWP1 and QWP2 (whose fast axis was fixed in the vertical direction). The output state became $|i\rangle = \cos(\vartheta_{\mathrm{Q}} - \pi/4)|\mathrm{D}\rangle + e^{-i2\vartheta_{\mathrm{Q}}}\sin(\vartheta_{\mathrm{Q}} - \pi/4)|\mathrm{A}\rangle$. The postselected state was fixed at $|\mathrm{f}\rangle = |\mathrm{H}\rangle$, as in case (i). The weak variance became the following purely imaginary number:

$$\sigma_{\mathrm{w}}^2(\hat{A}) = 2i\frac{\cos(2\vartheta_{\mathrm{Q}})}{\sin^2(2\vartheta_{\mathrm{Q}})}. \tag{9}$$

First, we observed the weak value in the transverse displacement of the beam's intensity distribution. Figure 4 plots the measured displacement of the mean of the beam's intensity distribution in the $\hat{X}$ basis, given by $\Delta\langle\hat{X}\rangle := \langle\hat{X}\rangle_{\mathrm{f}} - \langle\hat{X}\rangle_i$, as a function of ϑ_{H} in case (i). When ϑ_{H} is small, the pre- and postselected states are nearly orthogonal and $\Delta\langle\hat{X}\rangle$ becomes large. The theoretical curve of $\Delta\langle\hat{X}\rangle$ (blue solid curve in Fig. 4) was fitted to the measured values. The fitting parameters were the coupling strength θ, visibility V, and the technical error of the rotation angle of HWP $\Delta \in [-0.5°, 0.5°]$ (for details, see Appendix C). From

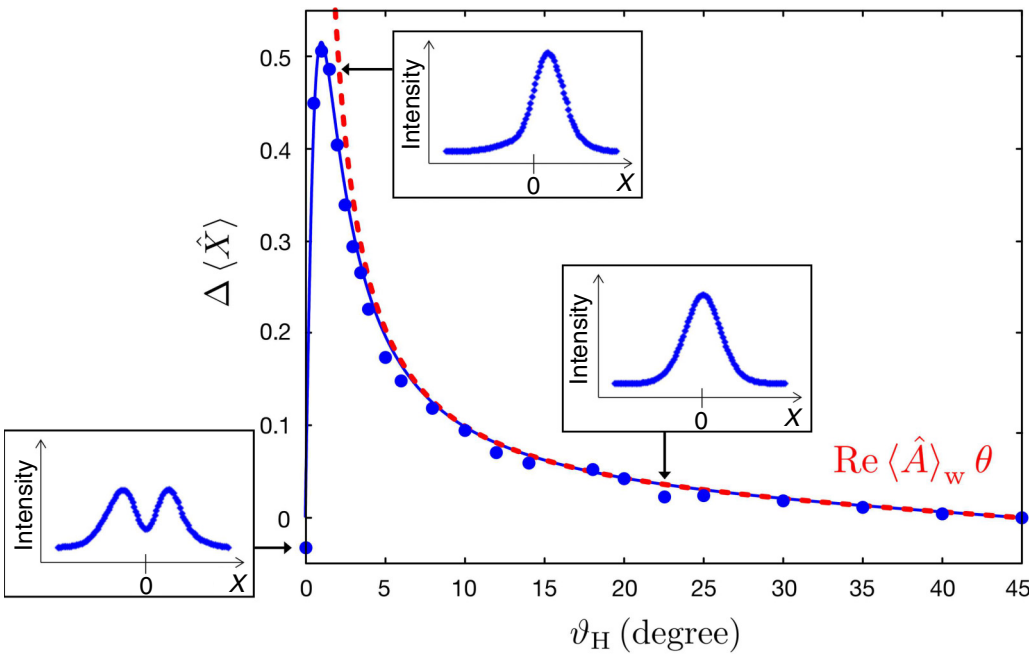

FIG. 4. Measurement results of $\Delta\langle\hat{X}\rangle$ in case (i) negative real weak variance. The solid blue curve is the theoretical fitting to the measured data (blue dots), and the red-dashed curve plots the theoretical weak value $\mathrm{Re}\langle\hat{A}\rangle_{\mathrm{w}}\theta$ ($\theta = 3.62 \times 10^{-2}$) vs ϑ_{H}. Insets show the intensity distributions of the beam for several values of ϑ_{H}. When ϑ_{H} is close to zero, the $O(\theta^3)$ term in Eq. (4) dominates and the distribution differs from a Gaussian one.

the fitting, we determined $\theta = 3.62 \times 10^{-2}$, $V = 1.00$, $\Delta = (1.57 \times 10^{-3})^{\circ}$. The weak value $\mathrm{Re}\langle\hat{A}\rangle_{\mathrm{w}}\theta$ ($\theta = 3.62 \times 10^{-2}$) versus ϑ_{H} is also plotted in Fig. 4 (red-dashed curve). Most of the measured values were consistent with the theoretical curve, verifying the measurements of the weak values.

We then observed the weak variance, which manifests as the changing variance of the beam's intensity distribution. Here $\Delta\sigma^2(\hat{M}) := [\sigma_{\mathrm{f}}^2(\hat{M}) - \sigma_i^2(\hat{M})]/\sigma_i^2(\hat{M})$ denotes the rate of variance change of the beam's intensity distribution from its initial value in the $\hat{M} = \hat{X}\cos\alpha + \hat{K}\sin\alpha$ basis. Figure 5 plots the measured $\Delta\sigma^2(\hat{M})$ ($\hat{M} = \hat{X}, \hat{\Omega}, \hat{K}, \hat{\Xi}$) as functions of ϑ_{H} and ϑ_{Q} in cases (i) and (ii), respectively. When ϑ_{H} and ϑ_{Q} were small, the pre- and postselected states were close to orthogonal and the variance changes were large. The theoretical curves of $\Delta\sigma^2(\hat{M})$ (blue solid curves in Fig. 5) were fitted to the measured values using θ, V, Δ, and the intensity of the background light N as fitting parameters (due to the visibility V being less than unity and the nonzero background noise N, $\Delta\sigma^2(\hat{M})$ differs slightly from the ideal weak variance even in the limit of $\theta \to 0$; see Appendix C for details). The theoretical weak variances $[\cos(2\alpha)\mathrm{Re}\sigma_{\mathrm{w}}^2(\hat{A}) + \sin(2\alpha)\mathrm{Im}\sigma_{\mathrm{w}}^2(\hat{A})]\theta^2$ (red-dashed curves) are also plotted as functions of ϑ_{H} or ϑ_{Q} in Fig. 5. Again, most of the measured values were consistent with the theoretical curves. In case (i), $\Delta\sigma^2(\hat{X})$ became negative because the weak variance was a negative real value; correspondingly, the $\Delta\sigma^2(\hat{K})$ increased. However, $\Delta\sigma^2(\hat{\Omega})$ and $\Delta\sigma^2(\hat{\Xi})$ remained almost zero because the imaginary part of the weak variance was zero. In case (ii), where the weak variance was positive and purely imaginary, $\Delta\sigma^2(\hat{\Omega})$ and $\Delta\sigma^2(\hat{\Xi})$ became positive and negative, respectively. However, $\Delta\sigma^2(\hat{X})$ and $\Delta\sigma^2(\hat{K})$ remained almost zero. Thus the real and imaginary parts of the weak variance appeared as width changes of the wave packet, in accordance with our theory.

IV. WEAK VARIANCE AS A STATISTIC OF THE WEAK-VALUED PROBABILITY DISTRIBUTION

In this section, we interpret the weak variance as a statistic of the weak-valued probability distribution, which is a pseudoprobability distribution in the pre- and postselected system. This relation is similar to that the variance is expressed as a statistic of the probability distribution in the preselected system. This statistical interpretation of the weak variance, together with the operational interpretation described above, rationalizes the definition of the weak variance as a counterpart of the variance in pre- and postselected systems.

We define the weak-valued probability $p_{\mathrm{w}j} := \langle\hat{\Pi}_j\rangle_{\mathrm{w}}$ as the weak value of each element of the set of projection operators $\{\hat{\Pi}_j\}_j$ that satisfy the completeness condition $\sum_j \hat{\Pi}_j = \hat{1}$ [Fig. 6]. Weak-valued probabilities can be any complex number outside [0,1], but their sum for all j is unity: $\sum_j p_{\mathrm{w}j} = 1$. By regarding the weak-valued probability as a quantity corresponding to the probability of finding a pre- and postselected particle in the eigenspace $\hat{\Pi}_j$ between the pre- and postselection, researchers have found a probabilistic approach to fundamental problems in quantum mechanics [4–10]. Because of their negativity and nonreality, weak-valued probabilities have played an essential role in studies such as the investigation of the relationship between disturbance and complementarity in quantum measurements [12,13], the explanation of the violation of Bell inequality using negative probabilities [11], quantum enhancement of the phase estimation sensitivity via postselection [24], and understanding

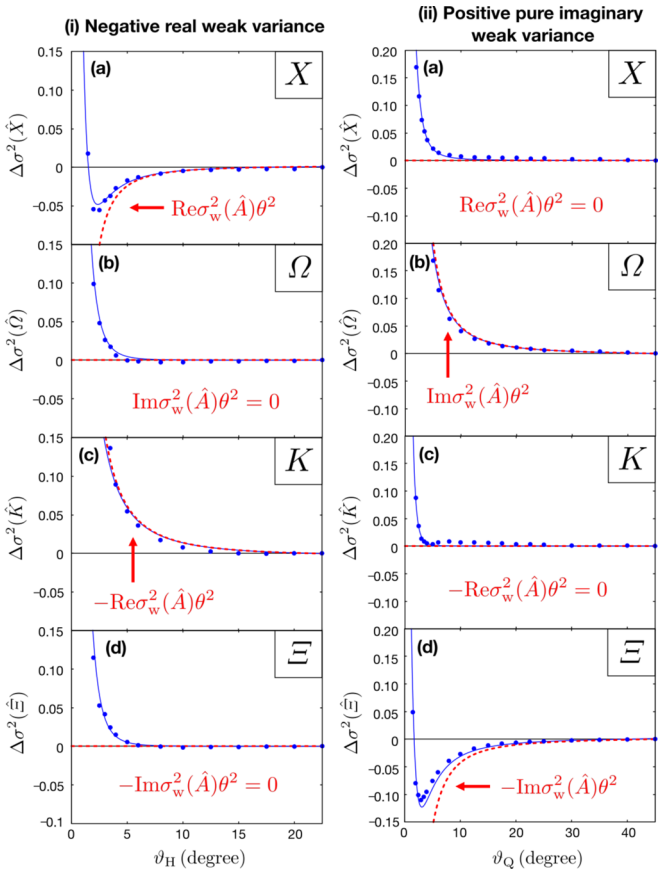

FIG. 5. Measurement results of $\Delta\sigma^2(\hat{M})$ in the cases of (i) negative real weak variance and (ii) purely imaginary weak variance. In (a), (b), (c), and (d), the measurement bases were $\hat{X}$, $\hat{\Omega}$, $\hat{K}$, and $\hat{\Xi}$, respectively. The solid blue lines are the theoretical fittings to the measured data (blue dots), and the red-dashed curves plot the theoretical weak variance $[\cos(2\alpha)\mathrm{Re}\sigma_{\mathrm{w}}^2(\hat{A}) + \sin(2\alpha)\mathrm{Im}\sigma_{\mathrm{w}}^2(\hat{A})]\theta^2$ vs ϑ_{H} or ϑ_{Q} (the θs were obtained by fitting the blue solid lines to the data). When ϑ_{H} and ϑ_{Q} approach zero, the $O(\theta^3)$ term in Eq. (7) dominates and the measured data notably deviate from the red-dashed lines.

out-of-time-order correlators as witnesses of quantum scrambling [25,26].

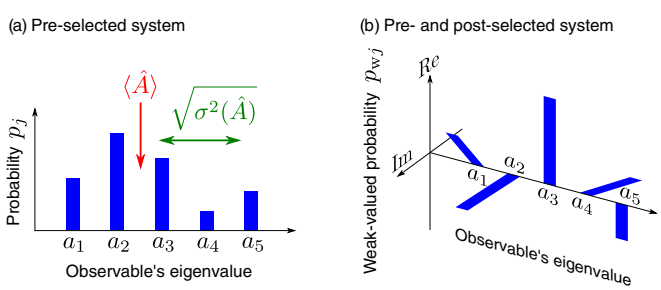

FIG. 6. (a) Probability distribution of the projective measurement of $\hat{A}$ in a preselected system, with an expectation value and variance of $\langle\hat{A}\rangle$ and $\sigma^2(\hat{A})$, respectively. (b) Weak-valued probability distribution of the observable $\hat{A}$ in pre- and postselected systems. The weak-valued probabilities $p_{\mathrm{w}j}$ are complex and expressed on the complex plane. $\langle\hat{A}\rangle_{\mathrm{w}}$ and $\sigma_{\mathrm{w}}^2(\hat{A})$ are also complex but cannot be depicted on this graph.

The weak value $\langle\hat{A}\rangle_{\mathrm{w}}$ and weak variance $\sigma_{\mathrm{w}}^2(\hat{A})$ of observable $\hat{A} = \sum_j a_j\hat{\Pi}_j$ can be expressed in terms of the weak-valued probabilities $\{p_{\mathrm{w}j}\}_j$ as follows:

$$\langle\hat{A}\rangle_{\mathrm{w}} = \sum_j a_j p_{\mathrm{w}j}, \tag{10}$$

$$\sigma_{\mathrm{w}}^2(\hat{A}) = \sum_j (a_j - \langle\hat{A}\rangle_{\mathrm{w}})^2 p_{\mathrm{w}j} = \sum_j |a_j - \langle\hat{A}\rangle_{\mathrm{w}}|^2 p_{\mathrm{w}j}, \tag{11}$$

where the second equation in Eq. (11) holds when $\hat{A}$ is Hermitian. These expressions are similar to those of the expectation value $\langle\hat{A}\rangle = \sum_j a_j p_j$ and variance $\sigma^2(\hat{A}) = \sum_j (a_j - \langle\hat{A}\rangle)^2 p_j$, respectively, obtained using probability distribution $\{p_j\}_j$. In this sense, the weak value and weak variance can be regarded as the expectation value and variance of a weak-valued probability distribution, respectively. In addition, as the weak-valued probability represents the conditional pseudoprobability of the Kirkwood-Dirac distribution [45,46], the weak value and weak variance are also regarded as the conditional pseudoexpectation value and conditional pseudovariance of the Kirkwood-Dirac distribution, respectively (see Appendix D for details). Furthermore, the weak value and weak variance satisfy the equations similar to the laws of total expectation and total variance, respectively:

$$\langle\hat{A}\rangle = \langle i|\hat{A}|i\rangle = \sum_j |\langle \mathrm{f}_j|i\rangle|^2 \langle\hat{A}\rangle_{\mathrm{w}j}, \tag{12}$$

$$\sigma^2(\hat{A}) = \sum_j |\langle \mathrm{f}_j|i\rangle|^2 \sigma_{\mathrm{w}j}^2(\hat{A}) + \sum_j |\langle \mathrm{f}_j|i\rangle|^2 \big(\langle\hat{A}\rangle_{\mathrm{w}j} - \langle\hat{A}\rangle\big)^2, \tag{13}$$

where $\langle\hat{A}\rangle_{\mathrm{w}j} := \langle \mathrm{f}_j|\hat{A}|i\rangle / \langle \mathrm{f}_j|i\rangle$ and $\sigma_{\mathrm{w}j}^2(\hat{A}) := \langle\hat{A}^2\rangle_{\mathrm{w}j} - \langle\hat{A}\rangle_{\mathrm{w}j}^2$.

Thus far, several definitions of the quantity corresponding to the variance in pre- and postselected systems have been considered. Examples are the weak variance expressed by Eq. (6) [27–31], its absolute value [32,33], its real part [34,39], and other forms [35–38]. The measurement method (indirect measurement) and statistical expression (a weak-valued probability distribution) of our proposed weak variance are similar to those of the conventional variance. Therefore the weak variance defined by Eq. (6) can be regarded as a reasonable counterpart of the variance in pre- and postselection systems.

V. CONCLUSION

We introduced the weak variance $\sigma_{\mathrm{w}}^2(\hat{A})$ as a complex counterpart of the variance in pre- and postselected systems. We theoretically showed that the weak variance appears as the changing width of the probe wave packet during indirect measurements of pre- and postselected systems and experimentally demonstrated the weak variance in an optical setup. We also expressed the weak value $\langle\hat{A}\rangle_{\mathrm{w}}$ and weak variance $\sigma_{\mathrm{w}}^2(\hat{A})$ as statistics of the weak-valued probability distribution $\{p_{\mathrm{w}j}\}_j$. These operational and statistical interpretations are similar to the expectation value $\langle\hat{A}\rangle$ and variance $\sigma^2(\hat{A})$ in preselected systems. Therefore our formulation of the weak

variance can be considered a reasonable definition of a counterpart of the variance in pre- and postselected systems.

Extending the concept of the weak variance, we then defined the *nth order weak moment* of the observable $\hat{A}$ as $\langle \hat{A}^n \rangle_{\rm w}$. The set of weak moments $\{\langle \hat{A}^k \rangle_{\rm w}\}_{k=1}^n$ fully characterizes the weak-valued probability distribution $\{p_{{\rm w}j}\}_{j=1}^n$. A similar relation exists between the set of moments $\{\langle \hat{A}^k \rangle\}_{k=1}^n$ and the probability distribution $\{p_j\}_{j=1}^n$. The nth order weak moment $\langle \hat{A}^n \rangle_{\rm w}$ can be experimentally observed in the indirect measurement setup by including terms up to the n-th order of θ in Eq. (3). Although the physical meanings of the weak moment is as elusive as the weak value, the weak moment may provide a new perspective on fundamental problems in quantum mechanics. For example, Scully *et al.*'s claim that the momentum disturbance associated with which-way measurement in Young's double-slit experiment can be avoided [47] has been justified by the negativity of the weak-valued probabilities corresponding to the momentum disturbance, which consequently have zero variance [12,13]. These studies are implicitly based on the weak variance (second-order weak moment) concept. Similarly, the weak moment is expected to play an important role in other problems of this type. In addition, measurement methods other than weak measurements with Gaussian probes—such as weak measurements using a qubit probe [48] and methods without a probe [49]—may find new implications for the weak moments.

Finally, as an application of the weak moments $\langle \hat{A}^n \rangle_{\rm w}$, we propose controlling the probe wave packet by pre- and postselection of the target system. In several studies, the probe wave packet was narrowed by appropriate pre- and postselection of the target system in the weak measurement setup [22,23,50]. If the higher order weak moments in the $O(\theta^2)$ term of Eq. (3) are properly controlled, we can configure any waveform of the probe (see Appendix E for details). For example, our method may represent a new construction method for the realization of the non-Gaussian states in the quadrature amplitude of light, such as the cat state [51,52] and the Gottesman-Kitaev-Preskill state [53], which play important roles in quantum optics.

ACKNOWLEDGMENTS

This research was supported by Japan Society for the Promotion of Science KAKENHI Grants No. 16K17524 and No. 19K14606, the Matsuo Foundation for Science Promotion, and the Research Foundation for Opto-Science and Technology.

APPENDIX A: CHANGE OF PROBE WAVE PACKET IN INDIRECT MEASUREMENTS OF MIXED PRE- AND POSTSELECTED SYSTEMS

We calculate the expectation value $\langle \hat{M} \rangle_{\rm f}$ and variance $\sigma_{\rm f}^2(\hat{M})$ of the probe wave packet in indirect measurements, when the pre- and postselected states are mixed states represented by density operators $\hat{\rho}_i$ and $\hat{\rho}_{\rm f}$, respectively. $\hat{\rho}_i = |i\rangle\langle i|$ and $\hat{\rho}_{\rm f} = |{\rm f}\rangle\langle {\rm f}|$ corresponds to the case of the pure pre- and postselected states, and $\hat{\rho}_{\rm f} = \hat{1}/d$ the case of the preselection alone. The time evolution of the entire state is calculated as

$$\begin{aligned}
\hat{\rho}_i \otimes |\phi\rangle\langle\phi| &\xrightarrow{\text{Interaction}} \exp(-i\theta \hat{A} \otimes \hat{K})(\hat{\rho}_i \otimes |\phi\rangle\langle\phi|)\exp(i\theta \hat{A} \otimes \hat{K}) \\
&\xrightarrow{\text{Post-selection}} {\rm tr}_{\rm t}\big[\hat{\rho}_{\rm f} \exp(-i\theta \hat{A} \otimes \hat{K})(\hat{\rho}_i \otimes |\phi\rangle\langle\phi|)\exp(i\theta \hat{A} \otimes \hat{K})\big] \\
&= {\rm tr}(\hat{\rho}_{\rm f}\hat{\rho}_i)\big[|\phi\rangle\langle\phi| + \big(-i\theta\langle\hat{A}\rangle_{\rm w}\hat{K}|\phi\rangle\langle\phi| - \tfrac{1}{2}\theta^2\langle\hat{A}^2\rangle_{\rm w}\hat{K}^2|\phi\rangle\langle\phi|\big) + {\rm H.c.} + \theta^2\tilde{A}\hat{K}|\phi\rangle\langle\phi|\hat{K}\big] + O(\theta^3) \\
&=: \tilde{\hat{\rho}}_\phi,
\end{aligned} \tag{A1}$$

where ${\rm tr}_{\rm t}$ denotes the partial trace in the target system and H.c. represents the Hermitian conjugate of the preceding term. $\langle\hat{A}\rangle_{\rm w} = {\rm tr}(\hat{\rho}_{\rm f}\hat{A}\hat{\rho}_i)/{\rm tr}(\hat{\rho}_{\rm f}\hat{\rho}_i)$ is the weak value of $\hat{A}$ in the pre- and postselected systems represented by density operators $\hat{\rho}_i$ and $\hat{\rho}_{\rm f}$, respectively. We now define $\tilde{A} := {\rm tr}(\hat{\rho}_{\rm f}\hat{A}\hat{\rho}_i\hat{A})/{\rm tr}(\hat{\rho}_{\rm f}\hat{\rho}_i)$. When $\hat{\rho}_i$ and $\hat{\rho}_{\rm f}$ are pure, $\tilde{A} = |\langle\hat{A}\rangle_{\rm w}|^2$, and when $\hat{\rho}_{\rm f} = \hat{1}/d$ (a completely mixed state), $\tilde{A} = {\rm tr}(\hat{\rho}_i\hat{A}^2) = \langle\hat{A}^2\rangle$.

The expectation value of $\hat{M}$ for the non-normalized probe state $\tilde{\hat{\rho}}_\phi$ is expressed as $\langle\hat{M}\rangle_{\rm f} = {\rm tr}(\tilde{\hat{\rho}}_\phi\hat{M})/{\rm tr}(\tilde{\hat{\rho}}_\phi)$. The numerator ${\rm tr}(\tilde{\hat{\rho}}_\phi\hat{M})$ is calculated as

$${\rm tr}(\tilde{\hat{\rho}}_\phi\hat{M}) = {\rm tr}(\hat{\rho}_{\rm f}\hat{\rho}_i)\big[\langle\hat{M}\rangle + \big(-i\theta\langle\hat{A}\rangle_{\rm w}\langle\hat{M}\hat{K}\rangle - \tfrac{1}{2}\theta^2\langle\hat{A}^2\rangle_{\rm w}\langle\hat{M}\hat{K}^2\rangle\big) + {\rm c.c.} + \theta^2\tilde{A}\langle\hat{K}\hat{M}\hat{K}\rangle\big] + O(\theta^3), \tag{A2}$$

where c.c. represents the complex conjugate of the preceding term. Because the expectation value of the product of odd numbers of $\hat{X}$ or $\hat{K}$ in our Gaussian probe state becomes zero, for $\hat{M} = \hat{X}\cos\alpha + \hat{K}\sin\alpha$, the above equation can be reduced to

$${\rm tr}(\tilde{\hat{\rho}}_\phi\hat{M}) = {\rm tr}(\hat{\rho}_{\rm f}\hat{\rho}_i)\big[\big(-i\theta\langle\hat{A}\rangle_{\rm w}\langle\hat{M}\hat{K}\rangle\big) + {\rm c.c.}\big] + O(\theta^3). \tag{A3}$$

By a similar calculation, the denominator ${\rm tr}(\tilde{\hat{\rho}}_\phi)$ is obtained as

$${\rm tr}(\tilde{\hat{\rho}}_\phi) = {\rm tr}(\hat{\rho}_{\rm f}\hat{\rho}_i)\big[1 + \big(-\tfrac{1}{2}\theta^2\langle\hat{A}^2\rangle_{\rm w}\langle\hat{K}^2\rangle\big) + {\rm c.c.} + \theta^2\tilde{A}\langle\hat{K}^2\rangle\big] + O(\theta^3). \tag{A4}$$

Therefore the expectation value $\langle\hat{M}\rangle_{\rm f}$ is expressed as

$$\langle\hat{M}\rangle_{\rm f} = \frac{{\rm tr}(\tilde{\hat{\rho}}_\phi\hat{M})}{{\rm tr}(\tilde{\hat{\rho}}_\phi)} = (-i\theta\langle\hat{A}\rangle_{\rm w}\langle\hat{M}\hat{K}\rangle) + {\rm c.c.} + O(\theta^3), \tag{A5}$$

where we have used the following formula:

$$\frac{a_0 + a_1\theta + a_2\theta^2 + O(\theta^3)}{b_0 + b_1\theta + b_2\theta^2 + O(\theta^3)} = \frac{a_0}{b_0} + \frac{a_1 b_0 - a_0 b_1}{b_0^2}\theta + \frac{a_2 b_0^2 - a_0 b_0 b_2 - a_1 b_0 b_1 + a_0 b_1^2}{b_0^3}\theta^2 + O(\theta^3). \tag{A6}$$

Because $\langle \hat{M}\hat{K}\rangle = (i\cos\alpha + \sin\alpha)/2$ in our Gaussian probe state, we obtain a concrete form of $\langle \hat{M}\rangle_{\rm f}$ as follows:

$$\langle \hat{M}\rangle_{\rm f} = \tfrac{1}{2}\theta\langle \hat{A}\rangle_{\rm w}(\cos\alpha - i\sin\alpha) + {\rm c.c.} + O(\theta^3) = \theta({\rm Re}\langle \hat{A}\rangle_{\rm w}\cos\alpha + {\rm Im}\langle \hat{A}\rangle_{\rm w}\sin\alpha) + O(\theta^3), \tag{A7}$$

which matches Eq. (4) in the main text.

The variance of $\hat{M}$ for the non-normalized probe state $\tilde{\hat{\rho}}_\phi$ is expressed as $\sigma_{\rm f}^2(\hat{M}) = \langle \hat{M}^2\rangle_{\rm f} - \langle \hat{M}\rangle_{\rm f}^2$. The first term is calculated as

$$\begin{aligned}
\langle \hat{M}^2\rangle_{\rm f} &= \frac{{\rm tr}(\tilde{\hat{\rho}}_\phi \hat{M}^2)}{{\rm tr}(\tilde{\hat{\rho}}_\phi)} \\
&= \frac{\langle \hat{M}^2\rangle + \left(-\frac{1}{2}\theta^2\langle \hat{A}^2\rangle_{\rm w}\langle \hat{M}^2\hat{K}^2\rangle\right) + {\rm c.c.} + \theta^2\tilde{A}\langle \hat{K}\hat{M}^2\hat{K}\rangle + O(\theta^3)}{1 + \left(-\frac{1}{2}\theta^2\langle \hat{A}^2\rangle_{\rm w}\langle \hat{K}^2\rangle\right) + {\rm c.c.} + \theta^2\tilde{A}\langle \hat{K}^2\rangle + O(\theta^3)} \\
&= \langle \hat{M}^2\rangle + \left(-\frac{1}{2}\theta^2\langle \hat{A}^2\rangle_{\rm w}\langle \hat{M}^2\hat{K}^2\rangle\right) + {\rm c.c.} + \theta^2\tilde{A}\langle \hat{K}\hat{M}^2\hat{K}\rangle + \langle \hat{M}^2\rangle\left(\frac{1}{2}\theta^2\langle \hat{A}^2\rangle_{\rm w}\langle \hat{K}^2\rangle + {\rm c.c.} - \theta^2\tilde{A}\langle \hat{K}^2\rangle\right) + O(\theta^3).
\end{aligned} \tag{A8}$$

Because the following equations hold for our Gaussian probe state:

$$\langle \hat{M}^2\rangle = \langle \hat{K}^2\rangle = \tfrac{1}{2}, \quad \langle \hat{M}^2\hat{K}\rangle = \tfrac{1}{4} - \tfrac{1}{2}[\cos(2\alpha) - i\sin(2\alpha)], \quad \text{and} \quad \langle \hat{K}\hat{M}^2\hat{K}\rangle = \tfrac{3}{4}, \tag{A9}$$

we obtain the following expression:

$$\langle \hat{M}^2\rangle_{\rm f} = \frac{{\rm tr}(\tilde{\hat{\rho}}_\phi \hat{M}^2)}{{\rm tr}(\tilde{\hat{\rho}}_\phi)} = \frac{1}{2} + \frac{1}{2}\theta^2\cos(2\alpha){\rm Re}\langle \hat{A}^2\rangle_{\rm w} + \frac{1}{2}\theta^2\sin(2\alpha){\rm Im}\langle \hat{A}^2\rangle_{\rm w} + \frac{1}{2}\theta^2\tilde{A} + O(\theta^3). \tag{A10}$$

The second term $\langle \hat{M}\rangle_{\rm f}^2$ is calculated using Eq. (A7) as

$$\langle \hat{M}\rangle_{\rm f}^2 = \tfrac{1}{2}\theta^2[({\rm Re}\langle \hat{A}\rangle_{\rm w})^2 - ({\rm Im}\langle \hat{A}\rangle_{\rm w})^2]\cos(2\alpha) + \theta^2{\rm Re}\langle \hat{A}\rangle_{\rm w}{\rm Im}\langle \hat{A}\rangle_{\rm w}\sin(2\alpha) + \tfrac{1}{2}\theta^2|\langle \hat{A}\rangle_{\rm w}|^2 + O(\theta^3). \tag{A11}$$

Therefore we obtain the concrete form of $\sigma_{\rm f}^2(\hat{M})$ as

$$\begin{aligned}
\sigma_{\rm f}^2(\hat{M}) &= \langle \hat{M}^2\rangle_{\rm f} - \langle \hat{M}\rangle_{\rm f}^2 \\
&= \tfrac{1}{2} + \tfrac{1}{2}\theta^2\cos(2\alpha)[{\rm Re}\langle \hat{A}^2\rangle_{\rm w} - ({\rm Re}\langle \hat{A}\rangle_{\rm w})^2 + ({\rm Im}\langle \hat{A}\rangle_{\rm w})^2] \\
&\quad + \tfrac{1}{2}\theta^2\sin(2\alpha)[{\rm Im}\langle \hat{A}^2\rangle_{\rm w} - 2{\rm Re}\langle \hat{A}\rangle_{\rm w}{\rm Im}\langle \hat{A}\rangle_{\rm w}] + \tfrac{1}{2}\theta^2(\tilde{A} - |\langle \hat{A}\rangle_{\rm w}|^2) + O(\theta^3) \\
&= \tfrac{1}{2} + \tfrac{1}{2}\theta^2\cos(2\alpha){\rm Re}\sigma_{\rm w}^2(\hat{A}) + \tfrac{1}{2}\theta^2\sin(2\alpha){\rm Im}\sigma_{\rm w}^2(\hat{A}) + \tfrac{1}{2}\theta^2(\tilde{A} - |\langle \hat{A}\rangle_{\rm w}|^2) + O(\theta^3),
\end{aligned} \tag{A12}$$

where we have used the following formulas:

$${\rm Re}\sigma_{\rm w}^2(\hat{A}) = {\rm Re}\langle \hat{A}^2\rangle_{\rm w} - ({\rm Re}\langle \hat{A}\rangle_{\rm w})^2 + ({\rm Im}\langle \hat{A}\rangle_{\rm w})^2, \tag{A13}$$

$${\rm Im}\sigma_{\rm w}^2(\hat{A}) = {\rm Im}\langle \hat{A}^2\rangle_{\rm w} - 2{\rm Re}\langle \hat{A}\rangle_{\rm w}{\rm Im}\langle \hat{A}\rangle_{\rm w}. \tag{A14}$$

In particular, when $\hat{\rho}_i$ and $\hat{\rho}_{\rm f}$ are pure, $\tilde{A} = |\langle \hat{A}\rangle_{\rm w}|^2$, so Eq. (A12) matches Eq. (7) in the main text. However, when $\hat{\rho}_{\rm f} = \hat{1}/d$ (preselection only), $\tilde{A} = \langle \hat{A}^2\rangle$, $\langle \hat{A}\rangle_{\rm w} = \langle \hat{A}\rangle$, and $\sigma_{\rm w}^2(\hat{A}) = \sigma^2(\hat{A}) \in \mathbb{R}$; therefore, we have

$$\sigma_{\rm f}^2(\hat{M}) = \tfrac{1}{2} + \tfrac{1}{2}\theta^2\cos(2\alpha)\sigma^2(\hat{A}) + \tfrac{1}{2}\theta^2\sigma^2(\hat{A}) + O(\theta^3). \tag{A15}$$

When $\alpha = 0$, we obtain

$$\sigma_{\rm f}^2(\hat{M}) = \sigma_{\rm f}^2(\hat{X}) = \tfrac{1}{2} + \theta^2\sigma^2(\hat{A}) + O(\theta^3), \tag{A16}$$

which equals $\sigma_{\rm f}^2(\hat{X})$ in Eq. (2) of the main text but without the $O(\theta^3)$ term (which vanishes in the full-order expansion in this case). Note that if the probe state is not a Gaussian wave packet, the expectation value and variance of $\hat{M}$ for the probe wave packet after the postselection will differ from Eqs. (A7) and (A12), respectively.

APPENDIX B: FRACTIONAL FOURIER TRANSFORM AND ITS OPTICAL REALIZATION

1. Definition of fractional Fourier transform

For any real number α, the α-angle fractional Fourier transform of a function ϕ is defined as

$$\mathcal{F}_\alpha[\phi](\omega) := \sqrt{\frac{1 - i\cot(\alpha)}{2\pi}} \int_{-\infty}^{\infty} dx\phi(x) \exp\left[i\left(\frac{\cot(\alpha)\omega^2 - 2\csc(\alpha)\omega x + \cot(\alpha)x^2}{2}\right)\right], \tag{B1}$$

where x and ω are dimensionless variables, $\cot(\alpha) = 1/\tan(\alpha)$, and $\csc(\alpha) = 1/\sin(\alpha)$. After transforming function ϕ_0 by $\mathcal{F}_\alpha$, we obtain a new function ϕ_α. When $\alpha = \pi/2$, $\mathcal{F}_\alpha$ reduces to the standard Fourier transform $\mathcal{F}$

$$\mathcal{F}_{\pi/2}[\phi_0](\omega) = \phi_{\pi/2}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\phi_0(x)e^{-i\omega x} = \mathcal{F}[\phi_0](\omega). \tag{B2}$$

The fractional Fourier transforms with $\alpha = \pm\pi/4$ and $\alpha = \pm 3\pi/4$ are respectively expressed as

$$\mathcal{F}_{\pm\pi/4}[\phi_0](\omega) = \phi_{\pm\pi/4}(\omega) = \sqrt{\frac{1 \mp i}{2\pi}} \int_{-\infty}^{\infty} dx\phi_0(x) \exp\left[\pm i\left(\frac{\omega^2 - 2\sqrt{2}\omega x + x^2}{2}\right)\right], \tag{B3}$$

$$\mathcal{F}_{\pm 3\pi/4}[\phi_0](\omega) = \phi_{\pm 3\pi/4}(\omega) = \sqrt{\frac{1 \pm i}{2\pi}} \int_{-\infty}^{\infty} dx\phi_0(x) \exp\left[\mp i\left(\frac{\omega^2 + 2\sqrt{2}\omega x + x^2}{2}\right)\right]. \tag{B4}$$

We call $\mathcal{F}_{\pm\pi/4}$ and $\mathcal{F}_{\pm 3\pi/4}$ the $\pm 1/2$- and $\pm 3/2$-Fourier transform, respectively.

2. Relationship between observables $\hat{X}$, $\hat{K}$, $\hat{\Omega}$, and $\hat{\Xi}$

The canonical conjugate of an observable $\hat{X}$, denoted by $\hat{K}$, satisfies the canonical commutation relation $[\hat{X}, \hat{K}] = i\hat{1}$. $\hat{X}$ and $\hat{K}$ are spectrally decomposed as follows:

$$\hat{X} = \int_{-\infty}^{\infty} dX X|X\rangle\langle X|, \quad \hat{K} = \int_{-\infty}^{\infty} dK K|K\rangle\langle K|. \tag{B5}$$

Their eigenvectors $|X\rangle$ and $|K\rangle$ are interrelated through the Fourier transform:

$$|K\rangle = \mathcal{F}_{-\pi/2}[|X\rangle](K) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dX|X\rangle e^{iKX}. \tag{B6}$$

Observables $\hat{\Omega}$ and $\hat{\Xi}$ are defined as

$$\hat{\Omega} := \frac{\hat{X} + \hat{K}}{\sqrt{2}}, \quad \hat{\Xi} := \frac{-\hat{X} + \hat{K}}{\sqrt{2}}. \tag{B7}$$

They satisfy the canonical commutation relation $[\hat{\Omega}, \hat{\Xi}] = i\hat{1}$. Observables $\hat{\Omega}$ and $\hat{\Xi}$ are spectrally decomposed as follows:

$$\hat{\Omega} = \int_{-\infty}^{\infty} d\Omega\Omega|\Omega\rangle\langle\Omega|, \quad \hat{\Xi} = \int_{-\infty}^{\infty} d\Xi\,\Xi|\Xi\rangle\langle\Xi|. \tag{B8}$$

The eigenvectors $|\Omega\rangle$ and $|\Xi\rangle$ of $\hat{\Omega}$ and $\hat{\Xi}$, respectively, are related to $|X\rangle$ by the $-1/2$- and $-3/2$-Fourier transforms, respectively:

$$|\Omega\rangle = \mathcal{F}_{-\pi/4}[|X\rangle](\Omega) = \sqrt{\frac{1+i}{2\pi}} \int_{-\infty}^{\infty} dX|X\rangle \exp\left[-i\left(\frac{\Omega^2 - 2\sqrt{2}\Omega X + X^2}{2}\right)\right], \tag{B9}$$

$$|\Xi\rangle = \mathcal{F}_{-3\pi/4}[|X\rangle](\Xi) = \sqrt{\frac{1-i}{2\pi}} \int_{-\infty}^{\infty} dX|X\rangle \exp\left[i\left(\frac{\Xi^2 + 2\sqrt{2}\Xi X + X^2}{2}\right)\right]. \tag{B10}$$

When state $|\phi\rangle$ is expanded in each basis as $|\phi\rangle = \int_{-\infty}^{\infty} dX\phi_0(X)|X\rangle = \int_{-\infty}^{\infty} d\Omega\phi_{\pi/4}(\Omega)|\Omega\rangle = \int_{-\infty}^{\infty} dK\phi_{\pi/2}(K)|K\rangle = \int_{-\infty}^{\infty} d\Xi\,\phi_{3\pi/4}(\Xi)|\Xi\rangle$, the relation between each wave function and basis vector is summarized as

$$\phi_0(X) \xrightarrow{\mathcal{F}_{\pi/4}} \phi_{\pi/4}(\Omega) \xrightarrow{\mathcal{F}_{\pi/4}} \phi_{\pi/2}(K) \xrightarrow{\mathcal{F}_{\pi/4}} \phi_{3\pi/4}(\Xi), \tag{B11}$$

$$|X\rangle \xrightarrow{\mathcal{F}_{-\pi/4}} |\Omega\rangle \xrightarrow{\mathcal{F}_{-\pi/4}} |K\rangle \xrightarrow{\mathcal{F}_{-\pi/4}} |\Xi\rangle. \tag{B12}$$

3. Optical realization of measurement of observables $\hat{X}$, $\hat{K}$, $\hat{\Omega}$, and $\hat{\Xi}$

This Appendix describes the optical system for measuring the observables $\hat{X}$, $\hat{K}$, $\hat{\Omega}$, and $\hat{\Xi}$ for a photon beam with a transverse distribution state $|\phi\rangle$. To measure the dimensionless transverse-position observable $\hat{X}$, we measure the photon's transverse position using a photon detector with suitable spatial resolution. To measure $\hat{K}$, we optically Fourier-transform

the photon's wavefunction $\phi_0(X) = \langle X|\phi\rangle$ and measure the transverse position of the resulting function $\phi_{\pi/2}(K)$. The optical Fourier transform is realized by combining a lens passage with free-space propagation. Similarly, $\hat{\Omega}$ and $\hat{\Xi}$ can be measured by optically 1/2- and 3/2-Fourier-transforming $\phi_0(X)$ and measuring the transverse positions of the resulting functions $\phi_{\pi/4}(\Omega)$ and $\phi_{3\pi/4}(\Xi)$, respectively. In what follows, we derive the optical 1/2- and 3/2-Fourier transform by combining the lens passage and free-space propagation.

We assume that the beam is propagating in the z direction and define x, k, and k_x as the transverse position, total wave number, and x component of the wave vector, respectively. We apply the paraxial approximation and assume that k does not depend on k_x because $k_x \ll k$. We then define the dimensionless variables $X := xk$ and $K_x := k_x/k$. Free-space propagation through distance d is represented in wave-number space by the following transfer function:

$$H_D^{\text{free}}(K_x) \propto \exp\left(-i\frac{DK_x^2}{2}\right), \tag{B13}$$

where $D := dk$ is the dimensionless distance. In position space, free-space propagation is represented by a convolution with the following function:

$$h_D^{\text{free}}(X) \propto \int_{-\infty}^{\infty} dK_x H_D^{\text{free}}(K_x) e^{iK_xX} \propto \exp\left(i\frac{X^2}{2D}\right). \tag{B14}$$

Meanwhile, passage through a lens with focal length f is represented in the position space by the following transfer function:

$$h_F^{\text{lens}}(X) \propto \exp\left(-i\frac{X^2}{2F}\right), \tag{B15}$$

where $F := fk$ is the dimensionless focal length. In wave-vector space, passage through this lens is represented by a convolution with the following function:

$$H_F^{\text{lens}}(K_x) \propto \int_{-\infty}^{\infty} dX h_F^{\text{lens}}(X) e^{-iK_xX} \propto \exp\left(i\frac{FK_x^2}{2}\right). \tag{B16}$$

If a photon with a transverse wave function $\phi_0(X)$ sequentially passes through a lens with focal length f, propagates in free space through distance d, and passes through another lens with focal length f, the resultant wave function is calculated as

$$\phi_0(X) \xrightarrow{\text{lens } f} \phi_0(X)\exp\left(-i\frac{X^2}{2F}\right) \tag{B17}$$

$$\xrightarrow{\text{free-space propagation } d} \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left(-i\frac{X'^2}{2F}\right)\exp\left[i\frac{(X-X')^2}{2D}\right] \tag{B18}$$

$$\xrightarrow{\text{lens } f} \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left(-i\frac{X'^2}{2F}\right)\exp\left[i\frac{(X-X')^2}{2D}\right]\exp\left(-i\frac{X^2}{2F}\right)$$

$$= \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left[i\frac{(F-D)X^2 - 2FXX' + (F-D)X'^2}{2FD}\right]. \tag{B19}$$

If we choose $D = F$, we obtain the standard Fourier transform of ϕ_0:

$$\text{Eq. } (B19) = \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left(i\frac{-XX'}{F}\right) \propto \mathcal{F}_{\pi/2}[\phi_0]\left(\frac{X}{F}\right) = \phi_{\pi/2}\left(\frac{X}{F}\right). \tag{B20}$$

After the Fourier transform, the scale of the wave function can be adjusted by adjusting the focal length F. If we choose $D = (1 - 1/\sqrt{2})F$, then

$$\begin{aligned}
\text{Eq. } (B19) &= \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left[i\frac{X^2 - 2\sqrt{2}XX' + X'^2}{(2-\sqrt{2})F}\right] \\
&= \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left\{\frac{i}{2}\left[\left(\frac{X}{\sqrt{(\sqrt{2}-1)F}}\right)^2 - 2\sqrt{2}\left(\frac{X}{\sqrt{(\sqrt{2}-1)F}}\right)\left(\frac{X'}{\sqrt{(\sqrt{2}-1)F}}\right) + \left(\frac{X'}{\sqrt{(\sqrt{2}-1)F}}\right)^2\right]\right\} \\
&\propto \mathcal{F}_{\pi/4}\left[\phi_{0,F}^-\right]\left(\frac{X}{\sqrt{(\sqrt{2}-1)F}}\right) \quad \left[\phi_{0,F}^-(X) := \phi_0\left(X\sqrt{(\sqrt{2}-1)F}\right)\right] \\
&= \phi_{\pi/4,F}^-\left(\frac{X}{\sqrt{(\sqrt{2}-1)F}}\right),
\end{aligned} \tag{B21}$$

where $\phi^{-}_{0,F}(X)$ is a scaled wave function of $\phi_0(X)$. In this manner, we obtain the 1/2-Fourier transform of $\phi_{F,0}(X)$. Similarly, if we choose $D = (1 + 1/\sqrt{2})F$, the 3/2-Fourier transform is obtained as follows:

$$\begin{aligned}
\text{Eq. }(B19) &= \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left[-i\frac{X^2 + 2\sqrt{2}XX' + X'^2}{(2+\sqrt{2})F}\right] \\
&= \int_{-\infty}^{\infty} dX'\phi_0(X')\exp\left\{-\frac{i}{2}\left[\left(\frac{X}{\sqrt{(\sqrt{2}+1)F}}\right)^2 + 2\sqrt{2}\left(\frac{X}{\sqrt{(\sqrt{2}+1)F}}\right)\left(\frac{X'}{\sqrt{(\sqrt{2}+1)F}}\right) + \left(\frac{X'}{\sqrt{(\sqrt{2}+1)F}}\right)^2\right]\right\} \\
&\propto \mathcal{F}_{3\pi/4}\left[\phi^{+}_{0,F}\right]\left(\frac{X}{\sqrt{(\sqrt{2}+1)F}}\right) \quad \left[\phi^{+}_{0,F}(X) := \phi_0\left(X\sqrt{(\sqrt{2}+1)F}\right)\right] \\
&= \phi^{-}_{3\pi/4,F}\left(\frac{X}{\sqrt{(\sqrt{2}+1)F}}\right),
\end{aligned} \tag{B22}$$

where $\phi^{+}_{0,F}(X)$ is a scaled wave function of $\phi_0(X)$.

Note that the second lens, which causes phase modulation in the position space, does not affect the measured intensity (projection) in the position basis. Therefore, in the experiment (see main text), the intensities of the beam's transverse distribution in the $\hat{X}$, $\hat{\Omega}$, $\hat{K}$ and $\hat{\Xi}$ bases were measured by inserting only one lens followed by free-space propagation.

APPENDIX C: THEORETICAL DERIVATIONS OF THE EXPECTATION VALUE AND VARIANCE IN OUR EXPERIMENTAL PROBE SYSTEM

First, we derive the exact formulas of the weak value and weak variance in our experimental setup. In the experiment, we assumed the preselected state $|i\rangle$, postselected state $|\mathrm{f}\rangle$, and the observable $\hat{A}$ as

$$|i\rangle = \cos\frac{\vartheta_i}{2}|0\rangle + e^{i\varphi_i}\sin\frac{\vartheta_i}{2}|1\rangle, \quad |\mathrm{f}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad \hat{A} = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{C1}$$

The weak value and weak variance are calculated as

$$\langle\hat{A}\rangle_{\mathrm{w}} = \frac{\cos\vartheta_i - i\sin\vartheta_i\sin\varphi_i}{1 + \sin\vartheta_i\cos\varphi_i}, \tag{C2}$$

$$\sigma^2_{\mathrm{w}}(\hat{A}) = \frac{2\sin\vartheta_i(\sin\vartheta_i + \cos\varphi_i) + i\sin(2\vartheta_i)\sin\varphi_i}{(1 + \sin\vartheta_i\cos\varphi_i)^2}. \tag{C3}$$

In our experiment, we used the following values:

$$\vartheta_i = \begin{cases} 4\vartheta_{\mathrm{H}} - \pi/2 & [\text{case (i)}] \\ 2\vartheta_{\mathrm{Q}} - \pi/2 & [\text{case (ii)}] \end{cases}, \quad \varphi_i = \begin{cases} 0 & [\text{case (i)}] \\ -2\vartheta_{\mathrm{Q}} & [\text{case (ii)}] \end{cases}. \tag{C4}$$

Substituting these terms into Eqs. (C2) and (C3), we obtain Eqs. (8) and (9), respectively.

Next, we derive the theoretical curves of the expectation value and variance of the probe wave packet demonstrated in our experiment. The state of the entire system after the interaction is

$$\exp(-i\theta\hat{A}\otimes\hat{K})|i\rangle|\phi\rangle = \cos\frac{\vartheta_i}{2}|0\rangle\exp(-i\theta\hat{K})|\phi\rangle + e^{i\varphi_i}\sin\frac{\vartheta_i}{2}|1\rangle\exp(i\theta\hat{K})|\phi\rangle. \tag{C5}$$

For notational simplicity, we denote the first and second terms on the right-hand side of Eq. (C5) by $|\Phi_0\rangle$ and $|\Phi_1\rangle$, respectively. Considering the decrease in visibility $V \in [0, 1]$, the state of the entire system after the interaction is expressed by the following density operator:

$$\hat{\rho} := |\Phi_0\rangle\langle\Phi_0| + |\Phi_1\rangle\langle\Phi_1| + V(|\Phi_0\rangle\langle\Phi_1| + |\Phi_1\rangle\langle\Phi_0|). \tag{C6}$$

After postselecting the target system onto $|\mathrm{f}\rangle$, the non-normalized probe state becomes $\tilde{\hat{\rho}}_{\mathrm{f}} := \langle\mathrm{f}|\hat{\rho}|\mathrm{f}\rangle$. The initial probe state is assumed as a Gaussian distribution $\langle X|\phi\rangle = \phi(X) = \pi^{-1/4}\exp(-X^2/2)$. The expectation value of the observable $\hat{M} = \hat{X}\cos\alpha + \hat{K}\sin\alpha$ for $\tilde{\hat{\rho}}_{\mathrm{f}}$ is calculated as

$$\langle\hat{M}\rangle_{\mathrm{f}} = \frac{\mathrm{tr}(\tilde{\hat{\rho}}_{\mathrm{f}}\hat{M})}{\mathrm{tr}(\tilde{\hat{\rho}}_{\mathrm{f}})} = \theta\frac{\cos\alpha\cos\vartheta_i + V\sin\alpha\sin\vartheta_i\sin\varphi_i e^{-\theta^2}}{1 + V\sin\vartheta_i\cos\varphi_i e^{-\theta^2}}. \tag{C7}$$

TABLE I. Results of fitting θ, V, Δ, and N to $\sigma_{\rm f}^2(\hat{M})$ ("-" denotes a fixed parameter).

	Case (i)				Case (ii)			
Measurement basis	θ	V	Δ	N	θ	V	Δ	N
$\hat{X}$	–	–	–	2.11×10^{-6}	5.56×10^{-2}	1.00	0.482°	8.21×10^{-7}
$\hat{\Omega}$	5.96×10^{-2}	1.00	0.185°	0.00	5.41×10^{-2}	1.00	-0.161°	0.00
$\hat{K}$	4.92×10^{-2}	1.00	0.500°	0.00	2.34×10^{-2}	0.999	-0.456°	0.00
$\hat{\Xi}$	5.24×10^{-2}	0.999	-0.108°	0.00	5.39×10^{-2}	0.999	0.407°	2.49×10^{-6}

To obtain the theoretical curve of the expectation value of $\hat{X}$ in case (i), we substitute Eq. (C4) and $\alpha = 0$ into Eq. (C7) as

$$\langle \hat{X} \rangle_{\rm f} = \theta \frac{\sin(4\vartheta_{\rm H})}{1 - V\cos(4\vartheta_{\rm H})e^{-\theta^2}}. \tag{C8}$$

Here we assume a technical error in the rotation angle of the HWP $\vartheta_{\rm H}$ $\Delta \in [-0.5^\circ, 0.5^\circ]$ that occurs in the experiment; accordingly, the fitting function is given as

$$\langle \hat{X} \rangle_{\rm f} = \theta \frac{\sin[4(\vartheta_{\rm H} + \Delta)]}{1 - V\cos[4(\vartheta_{\rm H} + \Delta)]e^{-\theta^2}}. \tag{C9}$$

Fitting this function to the measured data with θ, V and Δ as the fitting parameters, we obtained $\theta = 3.62 \times 10^{-2}$, $V = 1.00$, and $\Delta = 1.57 \times 10^{-3}$.

Meanwhile, the theoretical variance curve of the observable $\hat{M}$ for $\tilde{\hat{\rho}}_{\rm f}$, $\sigma_{\rm f}^2(\hat{M})$, is calculated as follows. The non-normalized probability density distribution of the signal $\tilde{\hat{\rho}}$ in the measurement basis $\hat{M} = \int_{-\infty}^{\infty} dM M|M\rangle\langle M|$ is given by $I_{\rm signal}(M) = \langle M|\tilde{\hat{\rho}}|M\rangle$. We now consider the effect of background light on the measured variances. The background light is modeled as the following rectangular function with intensity N and width $2L$:

$$I_{\rm background}(M) = \begin{cases} N & (M \in [-L, L]) \\ 0 & (\text{otherwise}), \end{cases} \tag{C10}$$

where $L = 5.6$ in our experimental setup. The normalized probability density distribution of the summed background and signal intensities is given by

$$I_{\rm total}(M) = \frac{I_{\rm signal}(M) + I_{\rm background}(M)}{\int_{-\infty}^{\infty} dM[I_{\rm signal}(M) + I_{\rm background}(M)]}. \tag{C11}$$

Using $I_{\rm total}(M)$, the theoretical curve of $\sigma_{\rm f}^2(\hat{M})$ is calculated as follows:

$$\begin{aligned} \sigma_{\rm f}^2(\hat{M}) &= \int_{-\infty}^{\infty} M^2 I_{\rm total}(M) - \left[\int_{-\infty}^{\infty} M I_{\rm total}(M)\right]^2 \\ &= \frac{1}{2} + \theta^2 \frac{\sin\vartheta_i\left\{\left(\cos^2\alpha - V^2\sin^2\alpha e^{-2\theta^2}\right)\sin\vartheta_i + Ve^{-\theta^2}[\cos(2\alpha)\cos\varphi_i + \sin(2\alpha)\cos\vartheta_i\sin\varphi_i]\right\}}{(1 + V\sin\vartheta_i\cos\varphi_i e^{-\theta^2} + 4LN)^2} \\ &\quad + 4LN\theta^2 \frac{\cos^2\alpha - Ve^{-\theta^2}\sin^2\alpha\sin\vartheta_i\cos\varphi_i}{(1 + V\sin\vartheta_i\cos\varphi_i e^{-\theta^2} + 4LN)^2} + \frac{2}{3}LN\frac{2L^2 - 3}{1 + V\sin\vartheta_i\cos\varphi_i e^{-\theta^2} + 4LN}. \end{aligned} \tag{C12}$$

The fitting functions in cases (i) and (ii) are derived by substituting Eq. (C4) into Eq. (C12) and replacing $\vartheta_{\rm H}$ and $\vartheta_{\rm Q}$ with $\vartheta_{\rm H} + \Delta$ and $\vartheta_{\rm Q} + \Delta$ ($\Delta \in [-0.5^\circ, 0.5^\circ]$), respectively. These functions were fitted to the measured data with θ, V, Δ, and N as the fitting parameters. The fitting results are summarized in Table I. When fitting the $\hat{X}$ measurement in case (i), N was the only fitting parameter. The other parameters were fixed at $\theta = 3.62 \times 10^{-2}$, $V = 1.00$, and $\Delta = (1.57 \times 10^{-3})^\circ$ because the experimental settings were unchanged from those of the $\langle \hat{X} \rangle_{\rm f}$ measurements.

APPENDIX D: WEAK VARIANCE AS A CONDITIONAL PSEUDOVARIANCE OF THE KIRKWOOD-DIRAC DISTRIBUTION

We show that the weak values and weak variances can be interpreted as conditional pseudoexpectation values and conditional pseudovariances of the Kirkwood-Dirac (KD) distribution [45,46], respectively. The (j, k) component of the KD distribution of state $|i\rangle$ can be expanded in two orthonormal bases $\{|a_j\rangle\}_j$ and $\{|a'_k\rangle\}_k$ as

$$D(a_j, a'_k|i) := \mathrm{tr}(|a_j\rangle\langle a_j|a'_k\rangle\langle a'_k|i\rangle\langle i|). \tag{D1}$$

The KD distribution is a joint pseudoprobability distribution representing the quantum state $|i\rangle$ and is generally a complex number. The KD distributions of states with indices j, k sum to unity: $\sum_{jk} D(a_j, a'_k|i) = 1$. The marginal distribution of the KD

distribution summed over one index becomes the projection probability distribution of $|i\rangle$ in the other basis:

$$\sum_j D(a_j, a'_k|i) = |\langle a'_k|i\rangle|^2, \quad \sum_k D(a_j, a'_k|i) = |\langle a_j|i\rangle|^2. \tag{D2}$$

When we choose $|a_j\rangle = |\mathrm{f}\rangle$, the conditional pseudoprobability $D(a'_k|i, \mathrm{f})$ of the KD distribution becomes

$$D(a'_k|i, \mathrm{f}) := \frac{D(\mathrm{f}, a'_k|i)}{\sum_k D(\mathrm{f}, a'_k|i)} = \frac{\langle \mathrm{f}|a'_k\rangle\langle a'_k|i\rangle}{\langle \mathrm{f}|i\rangle} = p'_{\mathrm{w}k}, \tag{D3}$$

where $\{p'_{\mathrm{w}k}\}_k$ is the weak-valued probability distribution of the pre- and postselection system $\{|i\rangle, |\mathrm{f}\rangle\}$ in the orthonormal basis $\{|a'_k\rangle\}_k$. Therefore the weak value $\langle \hat{A}'\rangle_{\mathrm{w}}$ and weak variance $\sigma^2_{\mathrm{w}}(\hat{A}')$ of observable $\hat{A}' := \sum_k a'_k|a'_k\rangle\langle a'_k|$ are represented as the conditional pseudoexpectation values and conditional pseudovariance of the KD distribution, respectively, as follows:

$$\sum_k a'_k D(a'_k|i, \mathrm{f}) = \sum_k a'_k p'_{\mathrm{w}k} = \langle \hat{A}'\rangle_{\mathrm{w}}, \tag{D4}$$

$$\sum_k (a'_k - \langle \hat{A}'\rangle_{\mathrm{w}})^2 D(a'_k|i, \mathrm{f}) = \sum_k (a'_k - \langle \hat{A}'\rangle_{\mathrm{w}})^2 p'_{\mathrm{w}k} = \sigma^2_{\mathrm{w}}(\hat{A}'). \tag{D5}$$

APPENDIX E: CONTROL OF THE PROBE WAVEFUNCTION BY PRE- AND POSTSELECTING THE TARGET SYSTEM

In indirect measurements of pre- and postselected systems, the probe state after the postselection can be controlled by appropriately choosing the pre- and postselected target system. The wave function $\tilde{\phi}(K)$ in the $\hat{K}$ basis of the probe state after the postselection $|\tilde{\phi}\rangle$ [Eq. (3) in the main text] is expressed for all orders of θ as

$$\tilde{\phi}(K) = \langle K|\tilde{\phi}\rangle = \langle K|\left[\langle \mathrm{f}|i\rangle \sum_{n=0}^{\infty} \frac{(-i\theta)^n}{n!} \langle \hat{A}^n\rangle_{\mathrm{w}} \hat{K}^n |\phi\rangle\right] = \langle \mathrm{f}|i\rangle \sum_{n=0}^{\infty} \frac{(-i\theta)^n}{n!} \langle \hat{A}^n\rangle_{\mathrm{w}} K^n \phi(K). \tag{E1}$$

Let $\phi_\star(K)$ be the wave function in the $\hat{K}$ basis of the desired probe state. To realize $\phi_\star(K)$ (except for a constant multiple), we can choose the weak moments $\{\langle \hat{A}^n\rangle_{\mathrm{w}}\}_n$ so that

$$\sum_{n=0}^{\infty} \frac{(-i\theta)^n}{n!} \langle \hat{A}^n\rangle_{\mathrm{w}} K^n \propto \frac{\phi_\star(K)}{\phi(K)}. \tag{E2}$$

When the target system is d-dimensional and $\hat{A}$ has full rank, we can independently choose the values of d weak moments $\langle \hat{A}^n\rangle_{\mathrm{w}}$. Therefore, by appropriately fixing the weak-moment values of the low-order terms of θ ($n = 1, \cdots, d$), which considerably affect the waveform, we can approximate the desired wave function.

[1] J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement* (Princeton University Press, 2014).

[2] J. von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition* (Princeton university press, 2018).

[3] 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).

[4] Y. Aharonov and L. Vaidman, Complete description of a quantum system at a given time, J. Phys. A **24**, 2315 (1991).

[5] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, Experimental realization of the quantum box problem, Phys. Lett. A **324**, 125 (2004).

[6] J. S. Lundeen and A. M. Steinberg, Experimental Joint Weak Measurement on a Photon Pair as a Probe of Hardy's Paradox, Phys. Rev. Lett. **102**, 020404 (2009).

[7] K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, Direct observation of hardy's paradox by joint weak measurement with an entangled photon pair, New J. Phys. **11**, 033011 (2009).

[8] 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**, 4492 (2014).

[9] R. Okamoto and S. Takeuchi, Experimental demonstration of a quantum shutter closing two slits simultaneously, Sci. Rep. **6**, 35161 (2016).

[10] A. Danan, D. Farfurnik, S. Bar-Ad, and L. Vaidman, Asking Photons Where They Have Been, Phys. Rev. Lett. **111**, 240402 (2013).

[11] B. L. Higgins, M. S. Palsson, G. Y. Xiang, H. M. Wiseman, and G. J. Pryde, Using weak values to experimentally determine "negative probabilities" in a two-photon state with bell correlations, Phys. Rev. A **91**, 012113 (2015).

[12] H. M. Wiseman, Directly observing momentum transfer in twin-slit "which-way" experiments, Phys. Lett. A **311**, 285 (2003).

[13] R. Mir, J. S. Lundeen, M. W. Mitchell, A. M. Steinberg, J. L. Garretson, and H. M. Wiseman, A double-slit 'which-way' experiment on the complementarity-uncertainty debate, New J. Phys. **9**, 287 (2007).

[14] Y. Xiao, H. M. Wiseman, J.-S. Xu, Y. Kedem, C.-F. Li, and G.-C. Guo, Observing momentum disturbance in double-slit "which-way" measurements, Sci. Adv. **5**, eaav9547 (2019).

[15] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, Experimental Joint Quantum Measurements with Minimum Uncertainty, Phys. Rev. Lett. **112**, 020401 (2014).

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

[17] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, Violation of Heisenberg'S Measurement-Disturbance Relationship by Weak Measurements, Phys. Rev. Lett. **109**, 100404 (2012).

[18] 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**, 1170 (2011).

[19] D. H. Mahler, L. Rozema, K. Fisher, L. Vermeyden, K. J. Resch, H. M. Wiseman, and A. M. Steinberg, Experimental nonlocal and surreal bohmian trajectories, Sci. Adv. **2**, e1501466 (2016).

[20] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan, Experimental Violation of Two-Party Leggett-Garg Inequalities with Semiweak Measurements Phys. Rev. Lett. **106**, 040402 (2011).

[21] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O'Brien, A. G. White, and G. J. Pryde, Violation of the leggett-garg inequality with weak measurements of photons, Proc. Natl. Acad. Sci. U.S.A. **108**, 1256 (2011).

[22] B. L. Bernardo, W. S. Martins, S. Azevedo, and A. Rosas, Uncertainty control and precision enhancement of weak measurements in the quadratic regime, Phys. Rev. A **92**, 012109 (2015).

[23] F. Matsuoka, A. Tomita, and Y. Shikano, Generation of phase-squeezed optical pulses with large coherent amplitudes by post-selection of single photon and weak cross-kerr non-linearity, Quantum Stud.: Math. Found. **4**, 159 (2017).

[24] D. R. M. Arvidsson-Shukur, N. Y. Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, Quantum negativity provides advantage in postselected metrology, Nat. Commun. **11**, 3775 (2020).

[25] N. Yunger Halpern, B. Swingle, and J. Dressel, Quasiprobability behind the out-of-time-ordered correlator, Phys. Rev. A **97**, 042105 (2018).

[26] J. R. González Alonso, N. Yunger Halpern, and J. Dressel, Out-of-Time-Ordered-Correlator Quasiprobabilities Robustly Witness Scrambling, Phys. Rev. Lett. **122**, 040404 (2019).

[27] B. Reznik and Y. Aharonov, Time-symmetric formulation of quantum mechanics, Phys. Rev. A **52**, 2538 (1995).

[28] B. Reznik, Interaction with a pre and post selected environment and recoherence, arXiv:quant-ph/9501023.

[29] A. Tanaka, Semiclassical theory of weak values, Phys. Lett. A **297**, 307 (2002).

[30] A. Brodutch, Weak measurements of non local variables, arXiv:0811.1706.

[31] A. D. Parks, Weak energy: Form and function, *Quantum Theory: A Two-Time Success Story* (Springer-Verlag, Mailand, 2014), pp. 291–302.

[32] Y. Aharonov and L. Vaidman, Properties of a quantum system during the time interval between two measurements, Phys. Rev. A **41**, 11 (1990).

[33] A. D. Parks, Weak covariance and the correlation of an observable with pre-selected and post-selected state energies during its time-dependent weak value measurement, Quantum Stud.: Math. Found. **5**, 455 (2018).

[34] M. R. Feyereisen, How the weak variance of momentum can turn out to be negative, Found. Phys. **45**, 535 (2015).

[35] H. F. Hofmann, Characterization of decoherence in a quantum channel using weak measurements, in *Proceedings of the International Quantum Electronics Conference and Conference on Lasers and Electro-Optics Pacific Rim 2011* (Optical Society of America, Sydney, 2011), p. I260.

[36] H. F. Hofmann, On the role of complex phases in the quantum statistics of weak measurements, New J. Phys. **13**, 103009 (2011).

[37] A. K. Pati and J. Wu, Uncertainty and complementarity relations in weak measurement, arXiv:1411.7218.

[38] Q.-C. Song and C.-F. Qiao, Uncertainty equalities and uncertainty relation in weak measurement, arXiv:1505.02233.

[39] P. P. Hofer, Quasi-probability distributions for observables in dynamic systems, Quantum **1**, 32 (2017).

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

[41] J. Dziewior, L. Knips, D. Farfurnik, K. Senkalla, N. Benshalom, J. Efroni, J. Meinecke, S. Bar-Ad, H. Weinfurter, and L. Vaidman, Universality of local weak interactions and its application for interferometric alignment, Proc. Natl. Acad. Sci. **116**, 2881 (2019).

[42] E. H. Kennard, Zur quantenmechanik einfacher bewegungstypen, Z. Phys. **44**, 326 (1927).

[43] H. P. Robertson, The uncertainty principle, Phys. Rev. **34**, 163 (1929).

[44] L. Vaidman, A. Ben-Israel, J. Dziewior, L. Knips, M. Weißl, J. Meinecke, C. Schwemmer, R. Ber, and H. Weinfurter, Weak value beyond conditional expectation value of the pointer readings, Phys. Rev. A **96**, 032114 (2017).

[45] J. G. Kirkwood, Quantum statistics of almost classical assemblies, Phys. Rev. **44**, 31 (1933).

[46] P. A. M. Dirac, On the analogy between classical and quantum mechanics, Rev. Mod. Phys. **17**, 195 (1945).

[47] M. O. Scully, B.-G. Englert, and H. Walther, Quantum optical tests of complementarity, Nature (London) **351**, 111 (1991).

[48] S-.J. Wu and K. Mølmer, Weak measurements with a qubit meter, Phys. Lett. A **374**, 34 (2009).

[49] K. Ogawa, H. Kobayashi, and A. Tomita, Operational formulation of weak values without probe systems, Phys. Rev. A **101**, 042117 (2020).

[50] A. D. Parks and J. E. Gray, Variance control in weak-value measurement pointers, Phys. Rev. A **84**, 012116 (2011).

[51] B. Yurke and D. Stoler, Generating Quantum Mechanical Superpositions of Macroscopically Distinguishable States via Amplitude Dispersion, Phys. Rev. Lett. **57**, 13 (1986).

[52] W. Schleich, M. Pernigo, and F. L. Kien, Nonclassical state from two pseudoclassical states, Phys. Rev. A **44**, 2172 (1991).

[53] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).