Significance of the imaginary part of the weak value

J. Dressel and A. N. Jordan

Phys. Rev. A **85**, 012107 — Published 23 January 2012

DOI: 10.1103/PhysRevA.85.012107

The significance of the imaginary part of the weak value

J. Dressel and A. N. Jordan
Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA
(Dated: December 16, 2011)

Unlike the real part of the generalized weak value of an observable, which can in a restricted sense be operationally interpreted as an idealized conditioned average of that observable in the limit of zero measurement disturbance, the imaginary part of the generalized weak value does not provide information pertaining to the observable being measured. What it does provide is direct information about how the initial state would be unitarily disturbed by the observable operator. Specifically, we provide an operational interpretation for the imaginary part of the generalized weak value as the logarithmic directional derivative of the post-selection probability along the unitary flow generated by the action of the observable operator. To obtain this interpretation, we revisit the standard von Neumann measurement protocol for obtaining the real and imaginary parts of the weak value and solve it exactly for arbitrary initial states and post-selections using the quantum operations formalism, which allows us to understand in detail how each part of the generalized weak value arises in the linear response regime. We also provide exact treatments of qubit measurements and Gaussian detectors as illustrative special cases, and show that the measurement disturbance from a Gaussian detector is purely decohering in the Lindblad sense, which allows the shifts for a Gaussian detector to be completely understood for any coupling strength in terms of a single complex weak value that involves the decohered initial state.

PACS numbers: 03.65.Ta,03.65.Ca,03.67.-a

I. INTRODUCTION

In their seminal Letter, Aharonov *et al.* [1] claimed that they could consistently assign a particular value to an observable that was being weakly measured in a pre- and post-selected ensemble. To illustrate their technique, they weakly coupled an observable $\hat{\boldsymbol{A}}$ to a continuous detector with an initial Gaussian wave-function. Normally, such a weak von Neumann coupling [2] would approximately shift the mean of the Gaussian detector wave-function by the expectation value $\langle\psi_i|\hat{\boldsymbol{A}}|\psi_i\rangle$ of $\hat{\boldsymbol{A}}$ in the initial state $|\psi_i\rangle$, which would effectively measure $\hat{\boldsymbol{A}}$; however, they showed that by post-selecting a final state $|\psi_f\rangle$ after the weak coupling, the mean of the Gaussian detector wave-function could be made to approximately shift by a *complex* quantity that they dubbed the *weak value* of the observable,

$$A_w = \frac{\langle\psi_f|\hat{\boldsymbol{A}}|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle}. \tag{1}$$

Notably, the weak value expression is not constrained to the eigenvalue range for the observable $\hat{\boldsymbol{A}}$, so it can become arbitrarily large for nearly orthogonal pre- and post-selections.

This complex shift in the mean of the Gaussian detector wave-function was only approximate under weak von Neumann coupling and not directly observable, so its significance was not overtly clear; however, the Letter [1] also showed that both the real and imaginary parts of (1) could be *operationally* obtained from the linear response of the detector under separate conjugate observable measurements. The practical benefit of this observation was that one could amplify the response of the detector by making a clever choice of post-selection, which potentially allowed for the sensitive determination of other small parameters contributing to the evolution.

After theoretical clarifications of the derivation in [3], experimental confirmation of such amplified detector response soon followed in optical systems [4, 5]. The amplification has since been used successfully to sensitively measure a variety of phenomena [6–11] to remarkable precision, using both the real and imaginary parts of (1) as amplification parameters. Several theoretical extensions of the original derivation of the amplification [12–27] and several proposals for other amplification measurements have also appeared [28–32]. In particular, it has been noted that how the amplification effect arises in such a continuous wave-function detector is not intrinsically quantum mechanical, but can also occur in classical wave mechanics [33], which has prompted recent study into the mathematical phenomenon of *superoscillations* (e.g. [34, 35]).

Conceptually, however, the weak value expression (1) has remained quite controversial: since it is generally complex and not constrained to the spectrum of $\hat{\boldsymbol{A}}$, how should it be interpreted? Its primary interpretation in the literature has rested somewhat loosely upon the observation that despite its anomalous behavior one can still decompose an expectation value through the insertion of the identity into an average of weak values, $\langle\psi_i|\hat{\boldsymbol{A}}|\psi_i\rangle = \sum_f |\langle\psi_f|\psi_i\rangle|^2(\langle\psi_f|\hat{\boldsymbol{A}}|\psi_i\rangle/\langle\psi_f|\psi_i\rangle)$, which has the same form as decomposing a classical expectation value $E(X|i)$ into an average of *conditioned expectation values* $E(X|i) = \sum_f P(f|i)E(X|i,f)$. This observation, together with its approximate appearance operationally in weak conditioned measurements, make it tempting to interpret the weak value as a disturbance-free counterfactual conditioned average that can be assigned to the observable within the context of a pre- and post-selected

ensemble even when it is not strictly measured [36–39].

Supporting this point of view is the fact that when the *real part* of (1) is bounded by the eigenvalue range of $\hat{\boldsymbol{A}}$, it agrees with the classical conditioned expectation value for the observable [37]. Moreover, even when the *real part* is outside the normal eigenvalue range, it still obeys a self-consistent logic [40] and seems to indicate oddly sensible information regarding the operator $\hat{\boldsymbol{A}}$. As such, it has been used quite successfully to analyze and interpret many quantum-mechanical paradoxes both theoretically and experimentally, such as tunneling time [41–44], vacuum Cherenkov radiation [45], cavity QED correlations [46], double-slit complementarity [47, 48], superluminal group velocities [49], the N-box paradox [50, 51], phase singularities [52], Hardy's paradox [53–56], photon arrival time [57], Bohmian trajectories [58–61], and Leggett-Garg inequality violations [62–64].

Arguably more important for its status as a quantity pertaining to the measurement of $\hat{\boldsymbol{A}}$, however, is the fact that the *real part* of (1) appears as a stable weak limit point for conditioned measurements even when the detector is not a von Neumann-coupled continuous wave that can experience superoscillatory interference (e.g. [63–67]). As a result, we can infer that at least the real part of (1) must have some operational significance specifically pertaining to the measurement of $\hat{\boldsymbol{A}}$ that extends beyond the scope of the original derivation. This observation prompted our Letter [68] showing that a principled treatment of a *general conditioned average* of an observable can in fact converge in the weak measurement limit to a generalized expression for the *real part* of (1),

$$\mathrm{Re}A_w = \frac{\mathrm{Tr}(\hat{\boldsymbol{P}}_f\{\hat{\boldsymbol{A}},\hat{\boldsymbol{\rho}}_i\})}{2\mathrm{Tr}(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i)}, \tag{2}$$

where $\{\hat{\boldsymbol{A}},\hat{\boldsymbol{\rho}}_i\} = \hat{\boldsymbol{A}}\hat{\boldsymbol{\rho}}_i + \hat{\boldsymbol{\rho}}_i\hat{\boldsymbol{A}}$ is the *anti-commutator* between the observable operator and an arbitrary initial state $\hat{\boldsymbol{\rho}}_i$ represented by a density operator, and where $\hat{\boldsymbol{P}}_f$ is an arbitrary post-selection represented by an element from a positive operator-valued measure (POVM). The general conditioned average converges to (2) provided that the manner in which $\hat{\boldsymbol{A}}$ is measured satisfies reasonable sufficiency conditions [69, 70] that ensure that the disturbance intrinsic to the measurement process does not persist in the weak limit.

It is in this precise restricted sense that we can operationally interpret the *real part* of the weak value (2) as an *idealized conditioned average of* $\hat{\boldsymbol{A}}$ *in the limit of zero measurement disturbance*. Since it is also the only apparent limiting value of the general conditioned average that no longer depends on *how* the measurement of $\hat{\boldsymbol{A}}$ is being made, it is also distinguished as a *measurement context-independent* conditioned average. These observations provide strong justification for the treatment of the real part of the weak value (2) as a form of value assignment [36–39, 71, 72] for the observable $\hat{\boldsymbol{A}}$ that depends only upon the preparation and post-selection [73].

However, we are still left with a mystery: what is the significance of the *imaginary part* of (1) that appears in the von Neumann measurement, and how does it relate to the operator $\hat{\boldsymbol{A}}$? We can find a partial answer to this question in existing literature (e.g. [12, 37, 41, 42]) that has associated the appearance of the imaginary part of (1) in the response of the detector with the intrinsic *disturbance*, or *back-action*, of the measurement process. For example, regarding continuous von Neumann detectors Aharonov and Botero [37, p.8] note that "the imaginary part of the complex weak value can be interpreted as a 'bias function' for the posterior sampling point [of the detector]." Furthermore, they note that "the weak value of an observable $\hat{\boldsymbol{A}}$ is tied to the role of $\hat{\boldsymbol{A}}$ as a generator for infinitesimal unitary transformations" [37, p.11]. Similarly, while discussing measurements of tunneling time Steinberg [41] states that the imaginary part is a "measure of the back-action on the particle due to the measurement interaction itself" and that the detector shift corresponding to the imaginary part "is sensitive to the details of the measurement apparatus (in particular, to the initial uncertainty in momentum), unlike the [shift corresponding to the real part]."

In this paper, we will augment these observations in the literature by providing a precise operational interpretation of the following generalized expression for the imaginary part of (1),

$$\mathrm{Im}A_w = \frac{\mathrm{Tr}(\hat{\boldsymbol{P}}_f(-i[\hat{\boldsymbol{A}},\hat{\boldsymbol{\rho}}_i]))}{2\mathrm{Tr}(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i)}, \tag{3}$$

where $[\hat{\boldsymbol{A}},\hat{\boldsymbol{\rho}}_i] = \hat{\boldsymbol{A}}\hat{\boldsymbol{\rho}}_i - \hat{\boldsymbol{\rho}}_i\hat{\boldsymbol{A}}$ is the *commutator* between $\hat{\boldsymbol{A}}$ and the initial state. We will see that the imaginary part of the weak value does not pertain to the measurement of $\hat{\boldsymbol{A}}$ as an observable. Instead, we will interpret it as half the *logarithmic directional derivative of the post-selection probability along the flow generated by the unitary action of the operator* $\hat{\boldsymbol{A}}$. As such, it provides an explicit measure for the idealized disturbance that the coupling to $\hat{\boldsymbol{A}}$ would have induced upon the initial state in the limit that the detector was not measured, which resembles the suggestion by Steinberg [41]; however, we shall see that the measurement of the detector can strongly alter the state evolution away from that ideal. The explicit commutator in (3) also indicates that the imaginary part of the weak value involves the operator $\hat{\boldsymbol{A}}$ in its role as a generator for unitary transformations as suggested by Aharonov and Botero [37], in contrast to the real part (2) that involves the operator $\hat{\boldsymbol{A}}$ in its role as a measurable observable.

To make it clear how the generalized weak value expressions (2) and (3) and their interpretations arise within a traditional von Neumann detector, we will provide an exact treatment of a von Neumann measurement using the formalism of quantum operations (e.g. [74–76]). In addition to augmenting existing derivations in the literature that are concerned largely with understanding the detector response (e.g. [12, 13, 15, 18–21, 23–27]), our exact approach serves to connect the standard treat-

ment of weak values to our more general contextual values analysis that produces the real part [68–70] more explicitly. We also provide several examples that specialize our exact solution to typically investigated cases: a particular momentum weak value, an arbitrary qubit observable measurement, and a Gaussian detector. As a consequence, we will show that the Gaussian detector is notable since it induces measurement disturbance that purely *decoheres* the system state into the eigenbasis of $\hat{\boldsymbol{A}}$ in the Lindblad sense with increasing measurement strength. Surprisingly, the pure decoherence allows the shifts in a Gaussian detector to be completely parametrized by a single complex weak value to all orders in the coupling strength, which allows those shifts to be completely understood using our interpretations of that weak value.

The paper is organized as follows. In §II we analyze the von Neumann measurement procedure in detail, starting with the traditional unconditioned analysis in §II A, followed by an operational analysis of the unconditioned case in §II B 1 and the conditioned case in §II B 2. After obtaining the exact solution for the von Neumann detector response, we consider the weak measurement regime to linear order in the coupling strength in §III, which clarifies the origins and interpretations of the expressions (2) and (3). We discuss the time-symmetric picture in §IV for completeness. After a brief Bohmian mechanics example in §V A that helps to illustrate our interpretation of the weak value, we provide the complete solutions for a qubit observable in §V B and a Gaussian detector in §V C. Finally, we present our conclusions in §VI.

II. VON NEUMANN MEASUREMENT

The traditional approach for obtaining a complex weak value [1] for a system observable is to post-select a weak Gaussian von Neumann measurement [2]. The real and imaginary parts of the complex weak value then appear as scaled shifts in the conditioned expectations of conjugate detector observables to linear order in the coupling strength. To clarify how these shifts occur and how the weak value can be interpreted, we shall solve the von Neumann measurement model exactly in the presence of post-selection.

A. Traditional Analysis

A von Neumann measurement [1, 2] unitarily couples an operator $\hat{\boldsymbol{A}}$ on a *system* Hilbert space $\mathcal{H}_s$ to a momentum operator $\hat{\boldsymbol{p}}$ on a continuous *detector* Hilbert space $\mathcal{H}_d$ via a time-dependent interaction Hamiltonian of the form,

$$\hat{\boldsymbol{H}}_I(t) = g(t)\hat{\boldsymbol{A}} \otimes \hat{\boldsymbol{p}}. \tag{4}$$

The interaction profile $g(t)$ is assumed to be a function that is only nonzero over some interaction time interval $t \in [0, T]$. The interaction is also assumed to be *impulsive* with respect to the natural evolution of the initial joint state $\hat{\boldsymbol{\rho}}$ of the system and detector; i.e., the interaction Hamiltonian (4) acts as the total Hamiltonian during the entire interaction time interval.

Solving the Schrödinger equation,

$$i\hbar\partial_t \hat{\boldsymbol{U}} = \hat{\boldsymbol{H}}_I \hat{\boldsymbol{U}}, \tag{5}$$

with the initial condition $\hat{\boldsymbol{U}}_0 = \hat{\mathbf{1}}$ produces a unitary operator,

$$\hat{\boldsymbol{U}}_T = \exp\left(\frac{g}{i\hbar}\hat{\boldsymbol{A}} \otimes \hat{\boldsymbol{p}}\right), \tag{6}$$

$$g = \int_0^T dt\, g(t), \tag{7}$$

that describes the full interaction over the time interval T. The constant g acts as an effective coupling parameter for the impulsive interaction. If the interaction is *weakly coupled* then g is sufficiently small so that $\hat{\boldsymbol{U}}_T \approx \hat{\mathbf{1}}$ and the effect of the interaction will be approximately negligible; however, we will make no assumptions about the weakness of the coupling *a priori*.

The unitary interaction (6) will entangle the system with the detector so that performing a direct measurement on the detector will lead to an *indirect* measurement being performed on the system. Specifically, we note that the position operator $\hat{\boldsymbol{x}}$ of the detector satisfies the canonical commutation relation $[\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}] = i\hbar\hat{\mathbf{1}}_d$, and thus will evolve in the Heisenberg picture of the interaction according to,

$$\begin{aligned}(\hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{x}})_T &= \hat{\boldsymbol{U}}_T^\dagger(\hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{x}})\hat{\boldsymbol{U}}_T, \\ &= \hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{x}} + g\hat{\boldsymbol{A}} \otimes \hat{\mathbf{1}}_d.\end{aligned} \tag{8}$$

As a result, measuring the mean of the detector position after the interaction $\langle x\rangle_T = \mathrm{Tr}((\hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{x}})_T\hat{\boldsymbol{\rho}})$ will produce,

$$\langle x\rangle_T = \langle x\rangle_0 + g\langle A\rangle_0. \tag{9}$$

Hence, the mean of the detector position will be shifted from its initial mean by the mean of the system observable $\hat{\boldsymbol{A}}$ in the *initial reduced system state*, linearly scaled by the coupling strength g. For this reason we say that directly measuring the average of the detector position $\hat{\boldsymbol{x}}$ results in an indirect measurement of the average of the system observable $\hat{\boldsymbol{A}}$.

The detector momentum $\hat{\boldsymbol{p}}$, on the other hand, does not evolve in the Heisenberg picture since $[\hat{\boldsymbol{U}}_T, \hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{p}}] = 0$. Hence, we expect that measuring the average detector momentum will provide no information about the system observable $\hat{\boldsymbol{A}}$.

As discussed in the introduction, however, when one conditions such a von Neumann measurement of the detector upon the outcome of a second measurement made only upon the system, then the conditioned average of *both* the position and the momentum of the detector can experience a shift. To see why this is so, we will find it useful to switch to the language of *quantum operations* (e.g. [74–76]) in order to dissect the measurement in more detail.

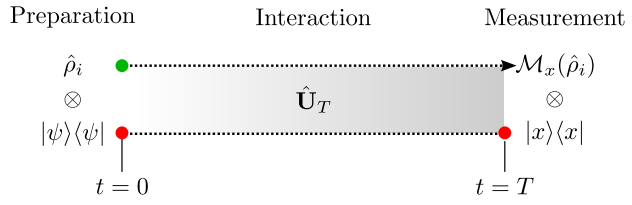

FIG. 1. (color online) Schematic for a von Neumann measurement. An initially prepared system state $\hat{\boldsymbol{\rho}}_i$ and detector state $|\psi\rangle\langle\psi|$ become entangled with the von Neumann unitary interaction $\hat{\boldsymbol{U}}_T$ (6) over a time interval T. Measuring a particular detector position x after the interaction updates the detector state to $|x\rangle\langle x|$ and also updates the system state to $\mathcal{M}_x(\hat{\boldsymbol{\rho}}_i)$, where $\mathcal{M}_x$ (11) is an effective measurement operation that encodes the entanglement with and subsequent measurement of the detector.

B. Quantum Operations

1. *Unconditioned Measurement*

As before, we will assume an impulsive interaction in what follows so that any natural time evolution in the joint system and detector state will be negligible on the time scale of the measurement. (For considerations of the detector dynamics, see [13].) We will also assume for simplicity of discussion that the initial joint state of the system and detector before the interaction is a product state and that the detector state is pure,

$$\hat{\boldsymbol{\rho}} = \hat{\boldsymbol{\rho}}_i \otimes |\psi\rangle\langle\psi|, \tag{10}$$

though we will be able to relax this assumption in our final results. Conceptually, this assumption states that a typical detector will be initially well-calibrated and uncorrelated with the unknown system state that is being probed via the interaction.

Evolving the initial state with the interaction unitary $\hat{\boldsymbol{U}}_T$ (6) will *entangle* the system with the detector. Hence, subsequently measuring a particular detector position will be equivalent to performing an *operation* $\mathcal{M}_x$ upon the reduced system state, as illustrated in Figure 1,

$$\mathcal{M}_x(\hat{\boldsymbol{\rho}}_i) = \mathrm{Tr}_d((\hat{\mathbf{1}}_s \otimes |x\rangle\langle x|)\hat{\boldsymbol{U}}_T\hat{\boldsymbol{\rho}}\hat{\boldsymbol{U}}_T^\dagger) = \hat{\boldsymbol{M}}_x\hat{\boldsymbol{\rho}}_i\hat{\boldsymbol{M}}_x^\dagger, \tag{11}$$

$$\hat{\boldsymbol{M}}_x = \langle x|\hat{\boldsymbol{U}}_T|\psi\rangle. \tag{12}$$

where $\mathrm{Tr}_d(\cdot)$ is the partial trace over the detector Hilbert space, and $\hat{\boldsymbol{M}}_x$ is the *Kraus operator* associated with the operation $\mathcal{M}_x$. Furthermore, since $\langle x|\psi\rangle = \psi(x)$ is the initial detector position wave-function we find $\hat{\boldsymbol{M}}_x = \int da\, \exp(-ga\partial_x)\psi(x)|a\rangle\langle a| = \int da\, \psi(x-ga)|a\rangle\langle a|$, or, more compactly, $\hat{\boldsymbol{M}}_x = \psi(x - g\hat{\boldsymbol{A}})$.

If we do *not* perform a subsequent post-selection on the system state, then we trace out the system to find the total probability density for detecting the position x,

$$p(x) = \mathrm{Tr}_s(\mathcal{M}_x(\hat{\boldsymbol{\rho}}_i)) = \mathrm{Tr}_s(\hat{\boldsymbol{E}}_x\hat{\boldsymbol{\rho}}_i), \tag{13}$$

$$\hat{\boldsymbol{E}}_x = \hat{\boldsymbol{M}}_x^\dagger\hat{\boldsymbol{M}}_x = \langle\psi|\hat{\boldsymbol{U}}_T^\dagger(\hat{\mathbf{1}}_s \otimes |x\rangle\langle x|)\hat{\boldsymbol{U}}_T|\psi\rangle, \tag{14}$$

where $\mathrm{Tr}_s(\cdot)$ is the partial trace over the system Hilbert space. The *probability operator* $\hat{\boldsymbol{E}}_x$ is a positive system operator that encodes the probability of measuring a particular detector position x, and can also be written in terms of the initial detector position wave-function as $\hat{\boldsymbol{E}}_x = |\psi(x - g\hat{\boldsymbol{A}})|^2$. To conserve probability it satisfies the condition, $\int dx\, \hat{\boldsymbol{E}}_x = \hat{\mathbf{1}}_s$, making the operators $\hat{\boldsymbol{E}}_x$ a *positive operator-valued measure* (POVM) on the system space.

Consequently, averaging the position of the detector will effectively average a system observable with the *initial system state*,

$$\langle x\rangle_T = \int_{-\infty}^{\infty} dx\, x\, p(x) = \mathrm{Tr}_s(\hat{\boldsymbol{O}}\hat{\boldsymbol{\rho}}_i), \tag{15}$$

$$\hat{\boldsymbol{O}} = \int_{-\infty}^{\infty} dx\, x\hat{\boldsymbol{E}}_x = \langle\psi|\hat{\boldsymbol{U}}_T^\dagger(\hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{x}})\hat{\boldsymbol{U}}_T|\psi\rangle,$$
$$= \langle x\rangle_0\hat{\mathbf{1}}_s + g\hat{\boldsymbol{A}},$$

where we see the Heisenberg evolved position operator (8) naturally appear.

Since the probability operators $\hat{\boldsymbol{E}}_x$ are diagonal in the basis of $\hat{\boldsymbol{A}}$, then the effective system operator $\hat{\boldsymbol{O}}$ will also be diagonal in the same basis. Hence, by modifying the values that we assign to the position measurements, we can arrange an indirect measurement of any system observable spanned by $\{\hat{\boldsymbol{E}}_x\}$ in the basis of $\hat{\boldsymbol{A}}$, including $\hat{\boldsymbol{A}}$ itself,

$$\hat{\boldsymbol{A}} = \int_{-\infty}^{\infty} dx\, \left(\frac{x - \langle x\rangle_0}{g}\right) \hat{\boldsymbol{E}}_x, \tag{16}$$

The chosen set of values $(x-\langle x\rangle_0)/g$ are *contextual values* for $\hat{\boldsymbol{A}}$, which can be thought of as a generalized spectrum that relates $\hat{\boldsymbol{A}}$ to the specific POVM $\{\hat{\boldsymbol{E}}_x\}$ associated with the *measurement context* $\{\mathcal{M}_x\}$ [68–70]. They are not the only values that we could assign to the position measurement in order to obtain the equality (16), but they are arguably the simplest to obtain and compute, as well as the most frequently used in the literature. It is in this precise sense that we can say that the von Neumann coupling leads to an indirect measurement of the average of $\hat{\boldsymbol{A}}$ in the absence of post-selection.

The measurement of $\hat{\boldsymbol{A}}$ comes at a cost, however, since the system state is necessarily *disturbed* by the operations $\mathcal{M}_x$ in order to obtain the probability operators $\hat{\boldsymbol{E}}_x$. The state may even be disturbed more than is strictly required to make the measurement of $\hat{\boldsymbol{A}}$, which can be seen by rewriting the measurement operators in polar form, $\hat{\boldsymbol{M}}_x = \hat{\boldsymbol{U}}_x|\hat{\boldsymbol{E}}_x|^{1/2}$, with the positive root of the probability operator $|\hat{\boldsymbol{E}}_x|^{1/2}$ and an additional unitary operator $\hat{\boldsymbol{U}}_x$. This decomposition implies that $\mathcal{M}_x$ splits into an

effective composition of two distinct operations,

$$\mathcal{M}_x(\hat{\boldsymbol{\rho}}_i) = \mathcal{U}_x(\mathcal{E}_x(\hat{\boldsymbol{\rho}}_i)), \tag{17a}$$
$$\mathcal{E}_x(\hat{\boldsymbol{\rho}}_i) = |\hat{\boldsymbol{E}}_x|^{1/2}\hat{\boldsymbol{\rho}}_i|\hat{\boldsymbol{E}}_x|^{1/2}, \tag{17b}$$
$$\mathcal{U}_x(\hat{\boldsymbol{\rho}}'_i) = \hat{\boldsymbol{U}}_x\hat{\boldsymbol{\rho}}'_i\hat{\boldsymbol{U}}_x^\dagger. \tag{17c}$$

We can interpret the operation $\mathcal{E}_x$ that involves only the roots of the probability operator $|\hat{\boldsymbol{E}}_x|^{1/2}$ as the *pure measurement operation* producing $\hat{\boldsymbol{E}}_x$. That is, it represents the *minimum necessary disturbance* that one must make to the initial state in order to extract a measurable probability. The second operation $\mathcal{U}_x$ unitarily disturbs the initial state, but does not contribute to $\hat{\boldsymbol{E}}_x$. Since only $\hat{\boldsymbol{E}}_x$ can be used to infer information about $\hat{\boldsymbol{A}}$ through the identity (16), we conclude that the disturbance from $\mathcal{U}_x$ is superfluous.

To identify the condition for eliminating $\mathcal{U}_x$, we can rewrite the Kraus operator (12) using the polar form of the initial detector position wave-function $\psi(x) = \exp(i\psi_s(x))\psi_r(x)$,

$$\hat{\boldsymbol{M}}_x = \exp(i\psi_s(x - g\hat{\boldsymbol{A}}))\psi_r(x - g\hat{\boldsymbol{A}}). \tag{18}$$

The phase factor becomes the unitary operator $\hat{\boldsymbol{U}}_x = \exp(i\psi_s(x - g\hat{\boldsymbol{A}}))$ for $\mathcal{U}_x$, while the magnitude becomes the required positive root $|\hat{\boldsymbol{E}}_x|^{1/2} = \psi_r(x - g\hat{\boldsymbol{A}})$ for $\mathcal{E}_x$. Hence, to eliminate the superfluous operation $\mathcal{U}_x$ from a von Neumann measurement with coupling Hamiltonian (4), one must use a *purely real* initial detector wave-function in position.

For contrast, measuring only a particular detector momentum p will be equivalent to performing a different operation $\mathcal{N}_p$ upon the reduced system state,

$$\mathcal{N}_p(\hat{\boldsymbol{\rho}}_i) = \mathrm{Tr}_d((\hat{\mathbf{1}}_s \otimes |p\rangle\langle p|)\hat{\boldsymbol{U}}_T\hat{\boldsymbol{\rho}}\hat{\boldsymbol{U}}_T^\dagger) = \hat{\boldsymbol{N}}_p\hat{\boldsymbol{\rho}}_i\hat{\boldsymbol{N}}_p^\dagger, \tag{19}$$
$$\hat{\boldsymbol{N}}_p = \langle p|\hat{\boldsymbol{U}}_T|\psi\rangle = \exp\left(\frac{gp}{i\hbar}\hat{\boldsymbol{A}}\right)\langle p|\psi\rangle. \tag{20}$$

The Kraus operator $\hat{\boldsymbol{N}}_p$ has a purely unitary factor containing $\hat{\boldsymbol{A}}$ that will disturb the system, regardless of the form of the initial momentum wave-function $\langle p|\psi\rangle$. Moreover, the probability operator associated with the momentum measurement has the form,

$$\hat{\boldsymbol{F}}_p = \hat{\boldsymbol{N}}_p^\dagger\hat{\boldsymbol{N}}_p = |\langle p|\psi\rangle|^2\hat{\mathbf{1}}_s, \tag{21}$$

which can only be used to measure the identity $\hat{\mathbf{1}}_s$.

For completeness we also briefly note that the conjugate Kraus operators $\hat{\boldsymbol{M}}_x$ and $\hat{\boldsymbol{N}}_p$ are related through a Fourier transform,

$$\hat{\boldsymbol{N}}_p = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\, e^{-ipx/\hbar}\hat{\boldsymbol{M}}_x, \tag{22a}$$
$$\hat{\boldsymbol{M}}_x = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dp\, e^{ipx/\hbar}\hat{\boldsymbol{N}}_p, \tag{22b}$$

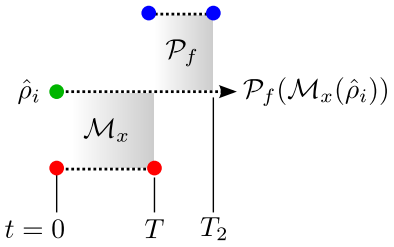

FIG. 2. (color online) Schematic for a sequence of two indirect measurements. After the von Neumann interaction and measurement of x illustrated in Figure 1 that produces the effective measurement operation $\mathcal{M}_x$ upon the initial system state, a second detector interacts impulsively with the system for a time interval $T_2 - T$. The second detector is then measured to have a particular outcome f, which updates the system state to $\mathcal{P}_f(\mathcal{M}_x(\hat{\boldsymbol{\rho}}_i))$, where $\mathcal{P}_f$ is another measurement operation. Taking the trace of the final system state will then produce the joint probability densities (24).

and that both detector probability operators can be obtained as marginals of a *Wigner quasi-probability operator* on the system Hilbert space,

$$\hat{\boldsymbol{W}}_{x,p} = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} dy\, e^{2ipy/\hbar}\hat{\boldsymbol{M}}_{x+y}^\dagger\hat{\boldsymbol{M}}_{x-y}, \tag{23a}$$
$$\hat{\boldsymbol{E}}_x = \int_{-\infty}^{\infty} dp\, \hat{\boldsymbol{W}}_{x,p}, \tag{23b}$$
$$\hat{\boldsymbol{F}}_p = \int_{-\infty}^{\infty} dx\, \hat{\boldsymbol{W}}_{x,p}. \tag{23c}$$

In the absence of interaction, then the Wigner quasi-probability operator reduces to the Wigner quasi-probability distribution $W(x,p)$ for the initial *detector* state, $\hat{\boldsymbol{W}}_{x,p} \xrightarrow{g=0} W(x,p)\hat{\mathbf{1}}_s$.

2. *Conditioned Measurement*

To post-select the system, an experimenter must perform a second measurement after the von Neumann measurement and filter the two-measurement event space based on the outcomes for the second measurement. In other words, the experimenter keeps only those pairs of outcomes for which the second outcome satisfies some constraint. The remaining measurement pairs can then be averaged to produce *conditioned averages* of the first measurement.

If we represent the second measurement as a set of probability operators $\{\hat{\boldsymbol{P}}_f\}$ indexed by some parameter f that can be derived analogously to (14) from a set of operations $\{\mathcal{P}_f\}$ as illustrated in Figure 2, then the total joint probability densities for the ordered sequences of measurement outcomes (x,f) and (p,f) will be,

$$p(x,f) = \mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{M}_x(\hat{\boldsymbol{\rho}}_i)) = \mathrm{Tr}_s(\hat{\boldsymbol{E}}_{x,f}\hat{\boldsymbol{\rho}}_i), \tag{24a}$$
$$p(p,f) = \mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{N}_p(\hat{\boldsymbol{\rho}}_i)) = \mathrm{Tr}_s(\hat{\boldsymbol{F}}_{p,f}\hat{\boldsymbol{\rho}}_i), \tag{24b}$$

where the joint probability operators,

$$\hat{\boldsymbol{E}}_{x,f} = \hat{\boldsymbol{M}}_x^\dagger \hat{\boldsymbol{P}}_f \hat{\boldsymbol{M}}_x, \tag{25a}$$

$$\hat{\boldsymbol{F}}_{p,f} = \hat{\boldsymbol{N}}_p^\dagger \hat{\boldsymbol{P}}_f \hat{\boldsymbol{N}}_p. \tag{25b}$$

are not simple products of the post-selection $\hat{\boldsymbol{P}}_f$ and the probability operators (14) or (21). Those operators can be recovered, however, by marginalizing over the index f, since the post-selection probability operators must satisfy a POVM condition $\sum_f \hat{\boldsymbol{P}}_f = \hat{\mathbf{1}}_s$.

The joint probabilities (24) will contain information not only about the first measurement and the initial system state, but also about the second measurement and any *disturbance* to the initial state that occurred due to the first measurement. In particular, the joint probability operators (25) can no longer satisfy the identity (16) due to the second measurement, so averaging the probabilities (24) must reveal more information about the measurement process than can be obtained solely from the operator $\hat{\boldsymbol{A}}$, the initial state $\hat{\boldsymbol{\rho}}_i$, and the post-selection $\hat{\boldsymbol{P}}_f$. As a poignant example, the unitary disturbance $\mathcal{U}_x$ in (17) that did not contribute to the operator identity (16) *will* contribute to the joint probability operators, $\hat{\boldsymbol{E}}_{x,f} = |\hat{\boldsymbol{E}}_x|^{1/2} \hat{\boldsymbol{U}}_x^\dagger \hat{\boldsymbol{P}}_f \hat{\boldsymbol{U}}_x |\hat{\boldsymbol{E}}_x|^{1/2}$.

The total probability for obtaining the post-selection outcome f can be obtained by marginalizing over either x or p in the joint probabilities,

$$\begin{aligned} p(f) &= \int_{-\infty}^{\infty} dx\, p(x,f) = \int_{-\infty}^{\infty} dp\, p(p,f), \qquad (26)\\ &= \mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{E}(\hat{\boldsymbol{\rho}}_i)),\\ \mathcal{E}(\hat{\boldsymbol{\rho}}_i) &= \mathrm{Tr}_d(\hat{\boldsymbol{U}}_T(\hat{\boldsymbol{\rho}}_i \otimes |\psi\rangle\langle\psi|)\hat{\boldsymbol{U}}_T^\dagger), \qquad (27) \end{aligned}$$

where the operation $\mathcal{E}$ is the total *non-selective* measurement that has been performed on $\hat{\boldsymbol{\rho}}_i$. Since $\mathcal{E}$ is not the identity operation, it represents the total *disturbance* intrinsic to the measurement process. It includes unitary evolution of the reduced system state due to the interaction Hamiltonian (4), as well as *decoherence* stemming from entanglement with the measured detector.

By experimentally filtering the event pairs to keep only a particular outcome f of the second measurement, an experimenter can obtain the conditional probabilities,

$$p(x|f) = \frac{p(x,f)}{p(f)} = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{M}_x(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{E}(\hat{\boldsymbol{\rho}}_i))}, \tag{28a}$$

$$p(p|f) = \frac{p(p,f)}{p(f)} = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{M}_p(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{E}(\hat{\boldsymbol{\rho}}_i))}, \tag{28b}$$

which can then be averaged to find the exact conditioned averages for the detector position and momentum,

$${}_f\langle x\rangle_T = \int_{-\infty}^{\infty} dx\, x\, p(x|f) = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{X}_T(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{E}(\hat{\boldsymbol{\rho}}_i))}, \tag{29a}$$

$${}_f\langle p\rangle_T = \int_{-\infty}^{\infty} dp\, p\, p(p|f) = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{P}_T(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f \mathcal{E}(\hat{\boldsymbol{\rho}}_i))}, \tag{29b}$$

where,

$$\mathcal{X}_T(\hat{\boldsymbol{\rho}}_i) = \mathrm{Tr}_d((\hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{x}})\hat{\boldsymbol{U}}_T(\hat{\boldsymbol{\rho}}_i \otimes |\psi\rangle\langle\psi|)\hat{\boldsymbol{U}}_T^\dagger), \tag{30a}$$

$$\mathcal{P}_T(\hat{\boldsymbol{\rho}}_i) = \mathrm{Tr}_d((\hat{\mathbf{1}}_s \otimes \hat{\boldsymbol{p}})\hat{\boldsymbol{U}}_T(\hat{\boldsymbol{\rho}}_i \otimes |\psi\rangle\langle\psi|)\hat{\boldsymbol{U}}_T^\dagger), \tag{30b}$$

are detector averaging operations that affect the system state before the measurement of the post-selection is performed. It is worth noting at this point that we can relax the assumption (10) made about the initial state in the exact operational expressions (27) and (30). Similarly, if different contextual values are used to average the conditional probabilities in (29), then corresponding detector observables with the same spectra will appear in the operations (30) in place of $\hat{\boldsymbol{x}}$ or $\hat{\boldsymbol{p}}$; for example, averaging the values $\alpha(x) = (x - \langle x\rangle_0)/g$ used in (16) will replace the detector observable $\hat{\boldsymbol{x}}$ in (30) with $\hat{\boldsymbol{\alpha}} = \int_{-\infty}^{\infty} dx\, \alpha(x)\, |x\rangle\langle x|$.

To better interpret (30), we bring the detector operators inside the unitary operators in (30) using the canonical commutation relations as in (8),

$$\mathcal{X}_T(\hat{\boldsymbol{\rho}}_i) = \mathcal{X}(\hat{\boldsymbol{\rho}}_i) + g\, \mathcal{E}(\{\hat{\boldsymbol{A}}, \hat{\boldsymbol{\rho}}_i\}/2), \tag{31a}$$

$$\mathcal{P}_T(\hat{\boldsymbol{\rho}}_i) = \mathcal{P}(\hat{\boldsymbol{\rho}}_i), \tag{31b}$$

which splits the $\mathcal{X}_T$ operation into two operations but only changes the form of $\mathcal{P}_T$. The operation proportional to g disturbs the symmetrized product $\{\hat{\boldsymbol{A}}, \hat{\boldsymbol{\rho}}_i\}/2 = (\hat{\boldsymbol{A}}\hat{\boldsymbol{\rho}}_i + \hat{\boldsymbol{\rho}}_i\hat{\boldsymbol{A}})/2$ of the *initial system state* with the operator $\hat{\boldsymbol{A}}$, while the operations,

$$\mathcal{X}(\hat{\boldsymbol{\rho}}_i) = \mathrm{Tr}_d(\hat{\boldsymbol{U}}_T(\hat{\boldsymbol{\rho}}_i \otimes \{\hat{\boldsymbol{x}}, |\psi\rangle\langle\psi|\}/2)\hat{\boldsymbol{U}}_T^\dagger), \tag{32a}$$

$$\mathcal{P}(\hat{\boldsymbol{\rho}}_i) = \mathrm{Tr}_d(\hat{\boldsymbol{U}}_T(\hat{\boldsymbol{\rho}}_i \otimes \{\hat{\boldsymbol{p}}, |\psi\rangle\langle\psi|\}/2)\hat{\boldsymbol{U}}_T^\dagger), \tag{32b}$$

disturb the symmetrized products of the *initial detector state* with the detector operators.

The form of the equations (31) clearly illustrates how the post-selection will affect the measurement. If the post-selection is the identity operator, $\hat{\boldsymbol{P}}_f = \hat{\mathbf{1}}_s$, then the unitary operators $\hat{\boldsymbol{U}}_T$ causing the total disturbance of the initial state will cancel through the cyclic property of the total trace in (29), leaving the averages in the *initial* states that were previously obtained,

$$\langle x\rangle_T = \langle x\rangle_0 + g\langle A\rangle_0, \tag{33a}$$

$$\langle p\rangle_T = \langle p\rangle_0. \tag{33b}$$

In this sense, commuting the detector operators $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{p}}$ in (30) through the unitary operators to arrive at (31) is equivalent to evolving them in the Heisenberg picture back from the time of measurement T to the initial time 0 in order to compare them with the initial states. However, the presence of the post-selection operator $\hat{\boldsymbol{P}}_f$ will now generally spoil the cancelation of the unitary operators that is implicit in the Heisenberg picture, leading to corrections from the disturbance between the pre- and post-selection.

The symmetrized products in (31) indicate the measurement being made on the *initial* states of the system and detector, which is then further disturbed by

the unitary operators $\hat{\boldsymbol{U}}_T$ as a consequence of the coupling Hamiltonian (4). The post-selection both conditions those measurements and reveals the disturbance, which corrects each term in (33), yielding the final exact expressions,

$$ {}_f\langle x\rangle_T = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{X}(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))} + g\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\{\hat{\boldsymbol{A}},\hat{\boldsymbol{\rho}}_i\}))}{2\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))}, \quad (34a) $$

$$ {}_f\langle p\rangle_T = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{P}(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))}. \quad (34b) $$

III. THE WEAK VALUE

If it were possible to leave the system state undisturbed while still allowing the measurement of $\hat{\boldsymbol{A}}$, then we would naïvely expect the disturbance $\mathcal{E}$ to reduce to the identity operation. Similarly, we would naïvely expect the operations $\mathcal{X}$ and $\mathcal{P}$ would reduce to $\langle x\rangle_0$ and $\langle p\rangle_0$ multiplying the identity operation, respectively. As a result, the conditioned averages (34) would differ from the unconditioned averages (33) solely by the replacement of the average $\langle A\rangle_0$ with the *real part* (2) of the complex *generalized weak value* expression,

$$ A_w = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{A}}\hat{\boldsymbol{\rho}}_i)}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i)}. \quad (35) $$

Since this expression depends solely upon the initial state $\hat{\boldsymbol{\rho}}_i$, the post-selection $\hat{\boldsymbol{P}}_f$, and the operator $\hat{\boldsymbol{A}}$, we are naïvely tempted to give $\mathrm{Re}A_w$ an intuitive interpretation as the *ideal conditioned expectation* of $\hat{\boldsymbol{A}}$ in a pre- and post-selected state with no intermediate measurement disturbance. However, it is strictly impossible to remove the disturbance from the measurement while still making the measurement, so we cannot rely on this sort of reasoning. We can make a similar interpretation in a restricted sense, however, by making the coupling strength g sufficiently small to reduce the disturbance to a minimal amount that still allows the measurement to be made.

To see how the operations $\mathcal{E}$, $\mathcal{X}$, and $\mathcal{P}$ in (27) and (32) depend on the coupling strength g, we expand them perturbatively,

$$ \mathcal{E}(\hat{\boldsymbol{\rho}}_i) = \sum_{n=0}\frac{1}{n!}\left(\frac{g}{i\hbar}\right)^n\langle p^n\rangle_0(\mathrm{ad}\hat{\boldsymbol{A}})^n(\hat{\boldsymbol{\rho}}_i), \quad (36a) $$

$$ \mathcal{X}(\hat{\boldsymbol{\rho}}_i) = \sum_{n=0}\frac{1}{n!}\left(\frac{g}{i\hbar}\right)^n\frac{\langle\{p^n,x\}\rangle_0}{2}(\mathrm{ad}\hat{\boldsymbol{A}})^n(\hat{\boldsymbol{\rho}}_i), \quad (36b) $$

$$ \mathcal{P}(\hat{\boldsymbol{\rho}}_i) = \sum_{n=0}\frac{1}{n!}\left(\frac{g}{i\hbar}\right)^n\langle p^{n+1}\rangle_0(\mathrm{ad}\hat{\boldsymbol{A}})^n(\hat{\boldsymbol{\rho}}_i), \quad (36c) $$

where the operation $(\mathrm{ad}\hat{\boldsymbol{A}})(\cdot) = [\hat{\boldsymbol{A}},\cdot]$ is the left action of $\hat{\boldsymbol{A}}$ in the adjoint representation of its Lie algebra, which takes the form of a commutator. That is, $(\mathrm{ad}\hat{\boldsymbol{A}})$ explicitly describes how $\hat{\boldsymbol{A}}$ *disturbs* the initial state due to the interaction that measures it.

The initial detector state plays a critical role in (36) by determining the various moments, $\langle p^n\rangle_0$, $\langle p^{n+1}\rangle_0$ and $\langle\{p^n,x\}/2\rangle_0$ that appear in the series expansions. Notably, if we make the initial detector wave-function purely real so that it minimally disturbs the system state then all moments containing odd powers of $\hat{\boldsymbol{p}}$ will vanish. We conclude that those moments of the disturbance operations are superfluous for obtaining the measurable probabilities that allow the measurement of $\hat{\boldsymbol{A}}$, while the moments with even powers of $\hat{\boldsymbol{p}}$ are necessary.

After expanding the corrections (34) to first order in g, we obtain the linear response of the conditioned detector means due to the interaction,

$$ {}_f\langle x\rangle_T \to \langle x\rangle_0 + \frac{g}{i\hbar}\frac{\langle\{p,x\}\rangle_0}{2}\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f(\mathrm{ad}\hat{\boldsymbol{A}})(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i)} + g\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\{\hat{\boldsymbol{A}},\hat{\boldsymbol{\rho}}_i\})}{2\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i)}, \quad (37a) $$

$$ {}_f\langle p\rangle_T \to \langle p\rangle_0 + \frac{g}{i\hbar}\langle p^2\rangle_0\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f(\mathrm{ad}\hat{\boldsymbol{A}})(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i)}. \quad (37b) $$

Measurements for which this linear response is a good approximation are known as *weak measurements*.

After introducing the complex generalized weak value (35), we can write the linear response formulas in a more compact form,

$$ {}_f\langle x\rangle_T = \langle x\rangle_0 + \frac{g}{\hbar}\frac{\langle\{p,x\}\rangle_0}{2}(2\mathrm{Im}A_w) + g\,\mathrm{Re}A_w, \quad (38a) $$

$$ {}_f\langle p\rangle_T = \langle p\rangle_0 + \frac{g}{\hbar}\langle p^2\rangle_0(2\mathrm{Im}A_w), \quad (38b) $$

in terms of not only its real part, but also *twice its imaginary part*.

If the initial detector position wave-function $\psi(x)$ is purely real, so that the measurement is minimally disturbing, then $\langle\{p,x\}/2\rangle_0$ will vanish, leaving only $\mathrm{Re}A_w$ in ${}_f\langle x\rangle_T$ as we naïvely reasoned before. However, the term proportional to $2\mathrm{Im}A_w$ will not vanish in ${}_f\langle p\rangle_T$ to linear order in g, making it an element of measurement disturbance that persists even for minimally disturbing weak measurements.

These linear response formulas for the von Neumann measurement have also been obtained and discussed in the literature with varying degrees of generality and rigor (e.g. [1, 3, 12, 13, 15, 18–20, 23–27, 40]). However, our derivation has a conceptual advantage in that we see explicitly how the origins of the real and imaginary parts of the weak value differ with respect to the measurement of $\hat{\boldsymbol{A}}$. We are therefore in a position to give concrete interpretations for each part.

The real part (2) of the weak value $\mathrm{Re}A_w$ stems directly from the part of the conditioned shift of the detector pointer that corresponds to the measurement of $\hat{\boldsymbol{A}}$ and does not contain any further perturbation induced by the measurement coupling that would be indicated by factors of $(\mathrm{ad}\hat{\boldsymbol{A}})$. As a result, it can be interpreted as an idealized limit point for the average of $\hat{\boldsymbol{A}}$ in the

initial state $\hat{\boldsymbol{\rho}}_i$ that has been conditioned on the post-selection $\hat{\boldsymbol{P}}_f$ without any appreciable intermediate measurement disturbance. To support this point of view, we have also shown in [68–70] that $\mathrm{Re}A_w$ appears naturally as such a limit point for minimally disturbing measurements that are not of von Neumann type, provided that those measurements satisfy reasonable sufficiency conditions regarding the measurability of $\hat{\boldsymbol{A}}$.

The imaginary part (3) of the weak value $\mathrm{Im}A_w$, on the other hand, stems directly from the *disturbance* of the measurement and explicitly contains $(\mathrm{ad}\hat{\boldsymbol{A}})$, which is the action of $\hat{\boldsymbol{A}}$ as a generator for unitary evolution due to the specific Hamiltonian (4). The factor $2\mathrm{Im}A_w$ appears in (38) along with information about the initial detector momentum that is being coupled to $\hat{\boldsymbol{A}}$ in the Hamiltonian (4), as well as factors of $\hbar$, in stark contrast to the real part. How then can it be interpreted?

The significance of $2\mathrm{Im}A_w$ becomes more clear once we identity the *directional derivative operation* that appears in its numerator,

$$\delta_A(\cdot) = -i(\mathrm{ad}\hat{\boldsymbol{A}})(\cdot). \tag{39}$$

That is, $\delta_A(\hat{\boldsymbol{\rho}}_i)$ indicates the *rate of change* of the initial state $\hat{\boldsymbol{\rho}}_i$ along a *flow* in state-space generated by $\hat{\boldsymbol{A}}$.

The directional derivative should be familiar from the Schrödinger equation written in the form $\partial_t\hat{\boldsymbol{\rho}} = [\hat{\boldsymbol{H}},\hat{\boldsymbol{\rho}}]/i\hbar = \delta_\Omega(\hat{\boldsymbol{\rho}})$, where the scaled Hamiltonian $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{H}}/\hbar$ is a characteristic frequency operator. The integration of this equation is a unitary operation in exponential form $\hat{\boldsymbol{\rho}}(t) = \exp(t\delta_\Omega)(\hat{\boldsymbol{\rho}}(0)) = \exp(-it\hat{\boldsymbol{\Omega}})\hat{\boldsymbol{\rho}}(0)\exp(it\hat{\boldsymbol{\Omega}})$ that specifies a *flow* in state space, which is a collection of curves that is parametrized both by a time parameter t and by the initial condition $\hat{\boldsymbol{\rho}}(0)$. Specifying the initial condition $\hat{\boldsymbol{\rho}}(0) = \hat{\boldsymbol{\rho}}_i$, picks out the specific curve from the flow that contains $\hat{\boldsymbol{\rho}}_i$. The directional derivative of the initial state along that specific curve is then defined in the standard way, $\partial_t\hat{\boldsymbol{\rho}}(t)|_{t=0} = \delta_\Omega(\hat{\boldsymbol{\rho}}_i)$.

The fact that the quantum state space is always a continuous manifold of states allows such a flow to be defined in a similar fashion using any Hermitian operator, such as $\hat{\boldsymbol{A}}$, as a generator. Analogously to time evolution, such a flow has the form of a unitary operation, $\hat{\boldsymbol{\rho}}(\varepsilon) = \exp(\varepsilon\delta_A)(\hat{\boldsymbol{\rho}}(0))$, where the real parameter ε for the flow has units inverse to $\hat{\boldsymbol{A}}$. Therefore, taking the directional derivative of $\hat{\boldsymbol{\rho}}_i$ along the specific curve of this flow that passes through $\hat{\boldsymbol{\rho}}_i$ will produce (39). For an explicit example that we will detail in §V B, the state-space of a qubit can be parametrized as the continuous volume of points inside the unit Bloch sphere; the derivative (39) produces the vector field illustrated in Figure 3 tangent to the flow corresponding to Rabi oscillations of the qubit.

With this intuition in mind, we define the post-selection probability for measuring $\hat{\boldsymbol{P}}_f$ given an initial state $\hat{\boldsymbol{\rho}}_i(\varepsilon) = \exp(\varepsilon\delta_A)(\hat{\boldsymbol{\rho}}_i)$ that is changing along the flow generated by $\hat{\boldsymbol{A}}$,

$$p_f(\varepsilon) = \mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i(\varepsilon)). \tag{40}$$

The logarithmic directional derivative of this post-selection probability then produces the factor $2\mathrm{Im}A_w$ that appears in (38),

$$2\mathrm{Im}A_w = \partial_\varepsilon \ln p_f(\varepsilon)\big|_{\varepsilon=0}, \tag{41}$$

which is our main result.

In words, *the imaginary part of the weak value is half the logarithmic directional derivative of the post-selection probability along the natural unitary flow generated by* $\hat{\boldsymbol{A}}$. It does not provide any information about the measurement of $\hat{\boldsymbol{A}}$ as an observable, but rather *indicates an instantaneous exponential rate of change in the post-selection probability due to disturbance of the initial state caused by* $\hat{\boldsymbol{A}}$ *in its role as a generator for unitary transformations*. Specifically, for small ε we have the approximate relation,

$$p_f(\varepsilon) \approx p_f(0)(1 + (2\mathrm{Im}A_w)\varepsilon). \tag{42}$$

For a pure initial state $\hat{\boldsymbol{\rho}}_i = |\psi_i\rangle\langle\psi_i|$ and a projective post-selection $\hat{\boldsymbol{P}}_f = |\psi_f\rangle\langle\psi_f|$, the expression (41) simplifies,

$$2\mathrm{Im}A_w = \partial_\varepsilon \ln|\langle\psi_f|\exp(-i\varepsilon\hat{\boldsymbol{A}})|\psi_i\rangle|^2\big|_{\varepsilon=0}. \tag{43}$$

Hence, the corrections containing $2\mathrm{Im}A_w$ that appear in (38) stem directly from how the specific von Neumann Hamiltonian (4) unitarily disturbs the initial system state infinitesimally prior to any additional disturbance induced by the measurement of the detector. Conceptually, the coupling induces a natural unitary flow of the initial system state generated by $\hat{\boldsymbol{A}}$, which for infinitesimal g changes the joint probability $p(p,f)$ for a specific p by the amount $(2\mathrm{Im}A_w)(gp/\hbar)$, where $gp/\hbar$ is the infinitesimal parameter ε in (41) that has units inverse to $\hat{\boldsymbol{A}}$. Averaging this correction to the joint probability with the detector observables $\hat{\boldsymbol{x}}$ or $\hat{\boldsymbol{p}}$ produces the correction terms in (38).

IV. TIME SYMMETRY

As noted in [40, 77], a quantum system that has been pre- and post-selected exhibits time symmetry. We can make the time symmetry more apparent in our operational treatment by introducing the *retrodictive state*,

$$\hat{\boldsymbol{\rho}}_f = \hat{\boldsymbol{P}}_f/\mathrm{Tr}(\hat{\boldsymbol{P}}_f), \tag{44}$$

associated with the post-selection (see, e.g., [78, 79]) and rewriting our main results in the time-reversed retrodictive picture.

After cancelling normalization factors, the detector response (34) for a system retrodictively prepared in the final state $\hat{\boldsymbol{\rho}}_f$ that has been conditioned on the *pre-selection* measurement producing the initial system state

$\hat{\boldsymbol{\rho}}_i$ has the form,

$$ {}_i\langle x\rangle_T = \frac{\mathrm{Tr}_s(\mathcal{X}^*(\hat{\boldsymbol{\rho}}_f)\hat{\boldsymbol{\rho}}_i)}{\mathrm{Tr}_s(\mathcal{E}^*(\hat{\boldsymbol{\rho}}_f)\hat{\boldsymbol{\rho}}_i)} + g\frac{\mathrm{Tr}_s(\{\mathcal{E}^*(\hat{\boldsymbol{\rho}}_f),\hat{\boldsymbol{A}}\}\hat{\boldsymbol{\rho}}_i)}{2\mathrm{Tr}_s(\mathcal{E}^*(\hat{\boldsymbol{\rho}}_f)\hat{\boldsymbol{\rho}}_i)}, \quad (45a) $$

$$ {}_i\langle p\rangle_T = \frac{\mathrm{Tr}_s(\mathcal{P}^*(\hat{\boldsymbol{\rho}}_f)\hat{\boldsymbol{\rho}}_i)}{\mathrm{Tr}_s(\mathcal{E}^*(\hat{\boldsymbol{\rho}}_f)\hat{\boldsymbol{\rho}}_i)}, \quad (45b) $$

where the retrodictive operations $\mathcal{E}^*$, $\mathcal{X}^*$, and $\mathcal{P}^*$ are the adjoints of the predictive operations in (27) and (32),

$$ \mathcal{E}^*(\hat{\boldsymbol{\rho}}_f) = \langle\psi|\hat{\boldsymbol{U}}_T^\dagger\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{U}}_T|\psi\rangle, \quad (46a) $$

$$ \mathcal{X}^*(\hat{\boldsymbol{\rho}}_f) = \langle\psi|\{\hat{\boldsymbol{x}},\hat{\boldsymbol{U}}_T^\dagger\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{U}}_T\}/2|\psi\rangle, \quad (46b) $$

$$ \mathcal{P}^*(\hat{\boldsymbol{\rho}}_f) = \langle\psi|\{\hat{\boldsymbol{p}},\hat{\boldsymbol{U}}_T^\dagger\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{U}}_T\}/2|\psi\rangle. \quad (46c) $$

Notably, the symmetric product with $\hat{\boldsymbol{A}}$ that appears in the detector response (45) involves the retrodictive state that has been evolved back to the initial time due to the non-selective measurement operation $\mathcal{E}^*$. Hence, in both pictures the measurement of $\hat{\boldsymbol{A}}$ is being made with respect to the same initial time.

After expanding the retrodictive operations perturbatively as in (36),

$$ \mathcal{E}^*(\hat{\boldsymbol{\rho}}_f) = \sum_{n=0}\frac{1}{n!}\left(\frac{g}{i\hbar}\right)^n\langle p^n\rangle_0(\mathrm{ad}^*\hat{\boldsymbol{A}})^n(\hat{\boldsymbol{\rho}}_f), \quad (47a) $$

$$ \mathcal{X}^*(\hat{\boldsymbol{\rho}}_f) = \sum_{n=0}\frac{1}{n!}\left(\frac{g}{i\hbar}\right)^n\frac{\langle\{p^n,x\}\rangle_0}{2}(\mathrm{ad}^*\hat{\boldsymbol{A}})^n(\hat{\boldsymbol{\rho}}_f), \quad (47b) $$

$$ \mathcal{P}^*(\hat{\boldsymbol{\rho}}_f) = \sum_{n=0}\frac{1}{n!}\left(\frac{g}{i\hbar}\right)^n\langle p^{n+1}\rangle_0(\mathrm{ad}^*\hat{\boldsymbol{A}})^n(\hat{\boldsymbol{\rho}}_f), \quad (47c) $$

where $(\mathrm{ad}^*\hat{\boldsymbol{A}})(\cdot) = -(\mathrm{ad}\hat{\boldsymbol{A}})(\cdot) = [\cdot,\hat{\boldsymbol{A}}]$ is the right action of $\hat{\boldsymbol{A}}$, then the linear response of the detector (38) can be written in terms of the retrodictive forms of the real and imaginary parts of the complex weak value,

$$ \mathrm{Re}A_w = \frac{\mathrm{Tr}_s(\{\hat{\boldsymbol{\rho}}_f,\hat{\boldsymbol{A}}\}\hat{\boldsymbol{\rho}}_i)}{2\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\rho}}_i)}, \quad (48) $$

$$ \begin{aligned} 2\mathrm{Im}A_w &= \frac{\mathrm{Tr}_s((-i[\hat{\boldsymbol{\rho}}_f,\hat{\boldsymbol{A}}])\hat{\boldsymbol{\rho}}_i)}{\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\rho}}_i)}, \quad (49)\\ &= \partial_\varepsilon\ln\mathrm{Tr}_s(\exp(-\varepsilon\delta_A)(\hat{\boldsymbol{\rho}}_f)\hat{\boldsymbol{\rho}}_i)\big|_{\varepsilon=0}, \end{aligned} $$

which should be compared with (2), (3), and (41).

We see that the imaginary part of the weak value can also be interpreted as half the logarithmic directional derivative of the pre-selection probability as the retrodictive state changes in the opposite direction along the flow generated by $\hat{\boldsymbol{A}}$.

V. EXAMPLES

A. Bohmian Mechanics

To make the preceding abstract discussion of the weak value more concrete, let us consider a special case that has been recently discussed by Leavens [58], Wiseman [59], and Hiley [61], where the operator $\hat{\boldsymbol{A}} = \hat{\boldsymbol{p}}$ being measured is the momentum operator of the system particle. Since the wave-number operator $\hat{\boldsymbol{k}} = -\hat{\boldsymbol{p}}/\hbar$ generates a flow that is parametrized by the position x, then we expect from the discussion surrounding (41) that the imaginary part of a momentum weak value will give information about how the post-selection probability will change along changes in position.

If we restrict our initial system state to be a pure state $\hat{\boldsymbol{\rho}}_i = |\phi\rangle\langle\phi|$, and post-select the measurement of the momentum on a particular position $\hat{\boldsymbol{P}}_f = |x\rangle\langle x|$, then the detector will have the linear response relations (38) with the complex weak value given by,

$$ p_w = \frac{\langle x|\hat{\boldsymbol{p}}|\phi\rangle}{\langle x|\phi\rangle} = \frac{-i\hbar\partial_x\phi(x)}{\phi(x)}. \quad (50) $$

We can split this value naturally into its real and imaginary parts by considering the polar decomposition of the initial system state $\phi(x) = r(x)\exp(iS(x))$,

$$ p_w = \hbar\partial_x S(x) - i\hbar\partial_x\ln r(x). \quad (51) $$

The real part of the weak value $\mathrm{Re}\,p_w = \hbar\partial_x S(x)$ is the phase gradient, or *Bohmian momentum* for the initial state, which we can now interpret operationally as the *average momentum conditioned on the subsequent measurement of a particular* x in the ideal limit of no measurement disturbance. This connection between the real part of a weak value and the Bohmian momentum that was pointed out in [58, 59] has recently allowed Kocsis *et al.* [60] to experimentally reconstruct the averaged Bohmian trajectories in an optical two-slit interference experiment using such a von Neumann measurement.

The imaginary part of the weak value, $\mathrm{Im}\,p_w = -\hbar\partial_x\ln r(x)$, on the other hand, is the logarithmic gradient of the root of the probability density $\rho(x) = |\phi(x)|^2 = r^2(x)$ for the particle at the point x. Written in the form,

$$ 2\mathrm{Im}\,p_w = -\hbar\partial_x\ln\rho(x), \quad (52) $$

it describes the instantanous exponential rate of positional change of the probability density with respect to the particular post-selection point x, as expected. This quantity, scaled by an inverse mass $1/m$, was introduced under the name "osmotic velocity" in the context of a stochastic interpretation of quantum mechanics developed by Nelson [80], where it produced a diffusion term in the stochastic equations of motion for a classical point particle with diffusion coefficient $\hbar/2m$. Nelson's interpretation was carefully contrasted with a stochastic interpretation for the Bohmian pilot wave by Bohm and Hiley [81], and the connection of the osmotic velocity with a weak value was recently emphasized by Hiley [61].

Hence, the imaginary part of the momentum weak value does not provide information about a measurement of the momentum in the initial state. Instead, it indicates the logarithmic directional derivative of the probability

density for measuring x along the flow generated by $\hat{\boldsymbol{p}}$. The scaled derivative $-\hbar\partial_x$ appears since $\hat{\boldsymbol{p}} = -\hbar\hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{k}}$ generates flow along the position x.

B. Qubit Observable

To make the full von Neumann measurement process more concrete, let us also consider a simple example where $\hat{\boldsymbol{A}}$ operates on the two-dimensional Hilbert space of a qubit. (See also [12, 13, 15, 18–21, 23–27].) We can in such a case simplify the perturbative expansions (36) using the following identities,

$$\hat{\boldsymbol{A}} = A\hat{\boldsymbol{\sigma}}_3, \tag{53a}$$

$$\hat{\boldsymbol{\rho}}_i = \frac{1}{2}\left(\hat{\mathbf{1}}_s + \sum_k r_k\hat{\boldsymbol{\sigma}}_k\right), \tag{53b}$$

$$[\hat{\boldsymbol{\sigma}}_j, \hat{\boldsymbol{\sigma}}_k] = 2i\epsilon_{jkl}\hat{\boldsymbol{\sigma}}_l, \tag{53c}$$

$$\{\hat{\boldsymbol{\sigma}}_j, \hat{\boldsymbol{\sigma}}_k\} = 2\delta_{jk}\hat{\mathbf{1}}_s, \tag{53d}$$

where $\{\hat{\boldsymbol{\sigma}}_k\}_{k=1}^3$ are the usual Pauli operators, the components of the initial system state $\{r_k\}_{k=1}^3$ are real and satisfy the inequality $0 \leq \sum_k r_k^2 \leq 1$, ϵ_{jkl} is the completely antisymmetric Levi-Civita pseudotensor, and δ_{jk} is the Kronecker delta. We have defined $\hat{\boldsymbol{\sigma}}_3$ to be diagonal in the eigenbasis of $\hat{\boldsymbol{A}}$ and have rescaled the spectrum of $\hat{\boldsymbol{A}}$ for simplicity to zero out its maximally mixed mean $\mathrm{Tr}_s(\hat{\boldsymbol{A}}\hat{\mathbf{1}}_s/2) = 0$. As a result, $\langle A\rangle_0 = Ar_3$.

It follows that for positive integer n the repeated actions of $\hat{\boldsymbol{A}}$ on the various qubit operators have the forms,

$$(\mathrm{ad}\hat{\boldsymbol{A}})^n(\hat{\mathbf{1}}_s) = 0, \tag{54a}$$

$$(\mathrm{ad}\hat{\boldsymbol{A}})^{2n-1}(\hat{\boldsymbol{\sigma}}_1) = i\hat{\boldsymbol{\sigma}}_2(2A)^{2n-1}, \tag{54b}$$

$$(\mathrm{ad}\hat{\boldsymbol{A}})^{2n}(\hat{\boldsymbol{\sigma}}_1) = \hat{\boldsymbol{\sigma}}_1(2A)^{2n}, \tag{54c}$$

$$(\mathrm{ad}\hat{\boldsymbol{A}})^{2n-1}(\hat{\boldsymbol{\sigma}}_2) = -i\hat{\boldsymbol{\sigma}}_1(2A)^{2n-1}, \tag{54d}$$

$$(\mathrm{ad}\hat{\boldsymbol{A}})^{2n}(\hat{\boldsymbol{\sigma}}_2) = \hat{\boldsymbol{\sigma}}_2(2A)^{2n}, \tag{54e}$$

$$(\mathrm{ad}\hat{\boldsymbol{A}})^n(\hat{\boldsymbol{\sigma}}_3) = 0, \tag{54f}$$

which collectively imply that,

$$(\mathrm{ad}\hat{\boldsymbol{A}})^{2n-1}(\hat{\boldsymbol{\rho}}_i) = \frac{i}{2}(2A)^{2n-1}\left(r_1\hat{\boldsymbol{\sigma}}_2 - r_2\hat{\boldsymbol{\sigma}}_1\right), \tag{55a}$$

$$(\mathrm{ad}\hat{\boldsymbol{A}})^{2n}(\hat{\boldsymbol{\rho}}_i) = \frac{1}{2}(2A)^{2n}\left(r_1\hat{\boldsymbol{\sigma}}_1 + r_2\hat{\boldsymbol{\sigma}}_2\right), \tag{55b}$$

and hence that the nonselective measurement operation has the exact form,

$$\mathcal{E}(\hat{\boldsymbol{\rho}}_i) = \hat{\boldsymbol{\rho}}_i + \frac{c(g)r_1 - s(g)r_2}{2}\hat{\boldsymbol{\sigma}}_1 + \frac{c(g)r_2 + s(g)r_1}{2}\hat{\boldsymbol{\sigma}}_2, \tag{56a}$$

$$c(g) = \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}\left(\frac{2Ag}{\hbar}\right)^{2n}\langle p^{2n}\rangle_0, \tag{56b}$$

$$s(g) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{(2n-1)!}\left(\frac{2Ag}{\hbar}\right)^{2n-1}\langle p^{2n-1}\rangle_0. \tag{56c}$$

The correction term can be interpreted as a Rabi oscillation of the qubit that has been perturbed by the coupling to the detector. Indeed, if the detector operator $\hat{\boldsymbol{p}}$ were replaced with a constant p, then the interaction Hamiltonian (4) would constitute an evolution term for the qubit that would induce Rabi oscillations around the $\hat{\boldsymbol{\sigma}}_3$ axis of the Bloch sphere, which would be the natural flow in state space generated by the action of $\hat{\boldsymbol{A}}$. With the substitution $\hat{\boldsymbol{p}} \to p$ then $\langle p^n\rangle_0 \to p^n$, so $c(g) \to \cos(2gAp/\hbar) - 1$ and $s(g) \to \sin(2gAp/\hbar)$, which restores the unperturbed Rabi oscillations.

Similarly, we find that the averaging operations for the detector position and momentum (32) have the exact forms,

$$\mathcal{X}(\hat{\boldsymbol{\rho}}_i) = \langle x\rangle_0\hat{\boldsymbol{\rho}}_i + \frac{c_x(g)r_1 - s_x(g)r_2}{2}\hat{\boldsymbol{\sigma}}_1 + \frac{c_x(g)r_2 + s_x(g)r_1}{2}\hat{\boldsymbol{\sigma}}_2, \tag{57a}$$

$$c_x(g) = \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}\left(\frac{2Ag}{\hbar}\right)^{2n}\frac{\langle\{p^{2n}, x\}\rangle_0}{2}, \tag{57b}$$

$$s_x(g) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{(2n-1)!}\left(\frac{2Ag}{\hbar}\right)^{2n-1}\frac{\langle\{p^{2n-1}, x\}\rangle_0}{2}, \tag{57c}$$

and,

$$\mathcal{P}(\hat{\boldsymbol{\rho}}_i) = \langle p\rangle_0\hat{\boldsymbol{\rho}}_i + \frac{c_p(g)r_1 - s_p(g)r_2}{2}\hat{\boldsymbol{\sigma}}_1 + \frac{c_p(g)r_2 + s_p(g)r_1}{2}\hat{\boldsymbol{\sigma}}_2, \tag{58a}$$

$$c_p(g) = \sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}\left(\frac{2Ag}{\hbar}\right)^{2n}\langle p^{2n+1}\rangle_0, \tag{58b}$$

$$s_p(g) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{(2n-1)!}\left(\frac{2Ag}{\hbar}\right)^{2n-1}\langle p^{2n}\rangle_0. \tag{58c}$$

These operations differ from $\mathcal{E}$ only in how the various moments of the initial detector distribution weight the series for the Rabi oscillation. In particular, given the substitutions $\hat{\boldsymbol{p}} \to p$ and $\hat{\boldsymbol{x}} \to x$, then $\langle\{p^n, x\}/2\rangle_0 \to p^n x$ and $\langle p^{n+1}\rangle_0 \to p^{n+1}$, so $c_x(g) \to x\,(\cos(2gAp/\hbar) - 1)$, $s_x(g) \to x\sin(2gAp/\hbar)$, $c_p(g) \to p\,(\cos(2gAp/\hbar) - 1)$, and $s_p(g) \to p\sin(2gAp/\hbar)$. Therefore, if the detector remained uncorrelated with the system the averaging operations would reduce to $\mathcal{X}(\hat{\boldsymbol{\rho}}_i) \to x\,\mathcal{E}(\hat{\boldsymbol{\rho}}_i)$ and

$\mathcal{P}(\hat{\boldsymbol{\rho}}_i) \to p\,\mathcal{E}(\hat{\boldsymbol{\rho}}_i)$, which are the decoupled intial detector means scaling the Rabi-oscillating qubit state.

Since we have assumed that $\hat{\boldsymbol{A}}$ does not have a component proportional to the identity, the symmetric product $\{\hat{\boldsymbol{A}}, \hat{\boldsymbol{\rho}}_i\}/2 = Ar_3\hat{\mathbf{1}}_s/2 = \langle A\rangle_0(\hat{\mathbf{1}}_s/2)$ for a qubit will act effectively as an inner product that extracts the part of the initial state proportional to $\hat{\boldsymbol{A}}$. Therefore, the correction to $\langle A\rangle_0$ in ${}_f\langle x\rangle_T$ that appears in (34a) has the simple form,

$$g\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\{\hat{\boldsymbol{A}}, \hat{\boldsymbol{\rho}}_i\}/2))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))} = \frac{g\langle A\rangle_0}{\tilde{p}(f)}. \tag{59}$$

where the conditioning factor,

$$\begin{aligned}\tilde{p}(f) &= 2\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i)), \\ &= 2\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\rho}}_i) \\ &\quad + (c(g)r_1 - s(g)r_2)\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\sigma_1) \\ &\quad + (c(g)r_2 + s(g)r_1)\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\sigma_2),\end{aligned} \tag{60}$$

is $(2/\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f))$ times the total probability of obtaining the post-selection. We have expressed $\tilde{p}(f)$ more compactly in terms of the retrodictive state (44) to show how the deviations from the initial state that are induced by $\hat{\boldsymbol{A}}$ become effectively averaged by the post-selection state. In the absence of post-selection, the retrodictive state will be maximally mixed $\hat{\boldsymbol{\rho}}_f = \hat{\mathbf{1}}_s/2$ and $\tilde{p}(f) \to 1$, recovering the unconditioned average $\langle A\rangle_0$.

The correction to the detector mean position $\langle x\rangle_0$ in (34a) can be expressed in a similar way,

$$\begin{aligned}\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{X}(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))} &= \frac{1}{\tilde{p}(f)}\Big(2\langle x\rangle_0\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\rho}}_i) \\ &\quad + (c_x(g)r_1 - s_x(g)r_2)\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_1) \\ &\quad + (c_x(g)r_2 + s_x(g)r_1)\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_2)\Big),\end{aligned} \tag{61}$$

as can the correction to the detector mean momentum $\langle p\rangle_0$ in (34b),

$$\begin{aligned}\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{P}(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))} &= \frac{1}{\tilde{p}(f)}\Big(2\langle p\rangle_0\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\rho}}_i) \\ &\quad + (c_p(g)r_1 - s_p(g)r_2)\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_1) \\ &\quad + (c_p(g)r_2 + s_p(g)r_1)\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_2)\Big).\end{aligned} \tag{62}$$

Expanding (34) using (59), (61), and (62) to linear order in g, we find the linear response (38) in terms of the real and imaginary parts of the qubit weak value,

$$\mathrm{Re}A_w = \frac{\langle A\rangle_0}{2\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\rho}}_i)}, \tag{63a}$$

$$2\mathrm{Im}A_w = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\delta_A(\hat{\boldsymbol{\rho}}_i))}{\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\rho}}_i)}, \tag{63b}$$

$$\delta_A(\hat{\boldsymbol{\rho}}_i) = A(-r_2\hat{\boldsymbol{\sigma}}_1 + r_1\hat{\boldsymbol{\sigma}}_2). \tag{63c}$$

As expected, the real part contains information regarding the measurement of $\hat{\boldsymbol{A}}$ as an observable in the initial state, conditioned by the disturbance-free overlap between the predictive and retrodictive states. The imaginary part, on the other hand, contains $\delta_A(\hat{\boldsymbol{\rho}}_i)$, which is a tangent vector field on the Bloch sphere—illustrated in Figure 3—that corresponds to an infinitesimal portion of the Rabi oscillation being generated by $\hat{\boldsymbol{A}}$. This tangent vector field contains only the components r_1 and r_2 from bases *orthogonal* to $\hat{\boldsymbol{A}}$ in the initial state $\hat{\boldsymbol{\rho}}_i$, so $2\mathrm{Im}A_w$ contains only the retrodictive averages of corrections to bases *orthogonal* to $\hat{\boldsymbol{A}}$, and thus contains no information about the measurement of $\hat{\boldsymbol{A}}$ as an observable. As discussed in (41), $2\mathrm{Im}A_w$ is the logarithmic rate of change of the post-selection probability along the vector field $\delta_A(\hat{\boldsymbol{\rho}}_i)$. Scaling it by a small factor with units inverse to A will produce a probability correction to linear order. In the absence of post-selection, then $\hat{\boldsymbol{\rho}}_f \to \hat{\mathbf{1}}_s/2$, $\mathrm{Re}A_w \to \langle A\rangle_0$, and $\mathrm{Im}A_w \to 0$.

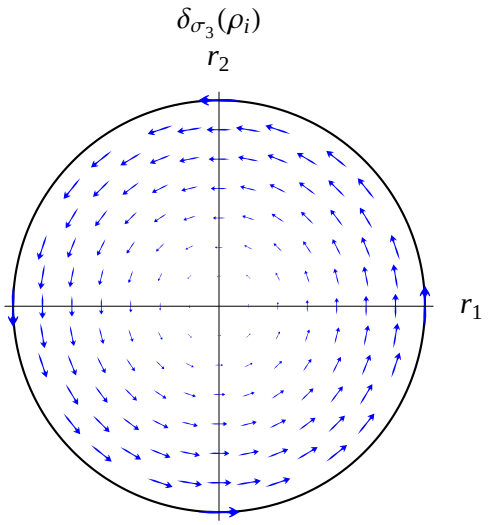

FIG. 3. (color online) The projection onto the plane $r_3 = 0$ of the qubit Bloch sphere, showing the vector field $\delta_{\sigma_3}(\hat{\boldsymbol{\rho}}_i) = -r_2\hat{\boldsymbol{\sigma}}_1 + r_1\hat{\boldsymbol{\sigma}}_2$ for arbitrary initial states $\hat{\boldsymbol{\rho}}_i = (\hat{\mathbf{1}} + r_1\hat{\boldsymbol{\sigma}}_1 + r_2\hat{\boldsymbol{\sigma}}_2 + r_3\hat{\boldsymbol{\sigma}}_3)/2$. The curves of the flow through this vector field are the Rabi oscillations around the r_3 axis that are generated by the unitary action of $\hat{\boldsymbol{\sigma}}_3$. The quantity $2\mathrm{Im}A_w$ (63) is the logarithmic rate of change of the post-selection probability (41) along this vector field.

C. Gaussian Detector

We can also apply our general results to the traditional case when the initial detector state in (10) is a zero-mean Gaussian in position,

$$\langle x|\psi\rangle = (2\pi\sigma^2)^{-1/4}\exp(-x^2/4\sigma^2), \tag{64}$$

Then the measurement operators for position detection (12) have the initial Gaussian form shifted by $g\hat{\boldsymbol{A}}$,

$$\hat{\boldsymbol{M}}_x = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-(x-g\hat{\boldsymbol{A}})^2/4\sigma^2}, \tag{65}$$

while the conjugate measurement operators for momentum detection (20) have the initial Gaussian modified by a unitary factor containing $\hat{\boldsymbol{A}}$,

$$\hat{\boldsymbol{N}}_p = \left(\frac{2\sigma^2}{\pi\hbar^2}\right)^{1/4} e^{-p^2\sigma^2/\hbar^2}\, e^{gp\hat{\boldsymbol{A}}/i\hbar}. \tag{66}$$

The Wigner quasi-probability operator (23a) correspondingly decouples into a product of Gaussian distributions, with only the position shifted by the system operator,

$$\hat{\boldsymbol{W}}_{x,p} = \frac{1}{\pi\hbar} e^{-(x-g\hat{\boldsymbol{A}})^2/2\sigma^2} e^{-2p^2\sigma^2/\hbar^2}. \tag{67}$$

Marginalizing the Wigner operator over momentum and position separately produces the probability operators (14) and (21),

$$\hat{\boldsymbol{E}}_x = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-g\hat{\boldsymbol{A}})^2/2\sigma^2}, \tag{68a}$$

$$\hat{\boldsymbol{F}}_p = \frac{\sigma}{\hbar}\sqrt{\frac{2}{\pi}}\, e^{-2p^2\sigma^2/\hbar^2}\, \hat{\mathbf{1}}_s. \tag{68b}$$

As anticipated, the probability operator for momentum no longer contains any information about the system operator $\hat{\boldsymbol{A}}$ and is proportional to the identity, so measuring the momentum provides zero information about any system operator not proportional to the identity.

In the presence of post-selection we can also exactly compute the disturbance operations (27), (30a), and (30b) using the following identities for the Gaussian detector moments,

$$\langle p^{2n}\rangle_0 = \left(\frac{\hbar}{2\sigma}\right)^{2n} (2n-1)!!, \tag{69a}$$

$$\langle p^{2n-1}\rangle_0 = 0, \tag{69b}$$

$$\langle \{p^n, x\}/2\rangle_0 = 0, \tag{69c}$$

$$\frac{(2n-1)!!}{(2n)!} = \frac{1}{2^n\, n!}, \tag{69d}$$

which hold for positive integer n. We find the simple results,

$$\mathcal{E}(\hat{\boldsymbol{\rho}}_i) = \exp\left(-\frac{1}{2}\left(\frac{g}{2\sigma}\right)^2 (\mathrm{ad}\hat{\boldsymbol{A}})^2\right)(\hat{\boldsymbol{\rho}}_i), \tag{70a}$$

$$\mathcal{X}(\hat{\boldsymbol{\rho}}_i) = 0, \tag{70b}$$

$$\mathcal{P}(\hat{\boldsymbol{\rho}}_i) = \frac{g}{i\hbar}\frac{\hbar^2}{4\sigma^2}(\mathrm{ad}\hat{\boldsymbol{A}})(\mathcal{E}(\hat{\boldsymbol{\rho}}_i)). \tag{70c}$$

The quantity $\epsilon = (g/2\sigma)^2$ with units inverse to $\hat{\boldsymbol{A}}^2$ emerges as the natural decoherence parameter, which we can see more clearly by rewriting the non-selective measurement operation in (70) as,

$$\hat{\boldsymbol{\rho}}_i(\epsilon) = \mathcal{E}(\hat{\boldsymbol{\rho}}_i) = \exp\left(\epsilon\mathcal{L}[\hat{\boldsymbol{A}}]\right)(\hat{\boldsymbol{\rho}}_i), \tag{71a}$$

$$\mathcal{L}[\hat{\boldsymbol{A}}](\hat{\boldsymbol{\rho}}_i) = \hat{\boldsymbol{A}}\hat{\boldsymbol{\rho}}_i\hat{\boldsymbol{A}}^\dagger - \frac{1}{2}\{\hat{\boldsymbol{\rho}}_i, \hat{\boldsymbol{A}}^\dagger\hat{\boldsymbol{A}}\}. \tag{71b}$$

The operation $\mathcal{L}[\hat{\boldsymbol{A}}](\hat{\boldsymbol{\rho}}_i)$ is the *Lindblad operation* [75, 76, 82] that produces *decoherence* in continuous dynamical systems, with $\hat{\boldsymbol{A}}$ playing the role of the Lindblad operator that decoheres the system. Since $\partial_\epsilon\hat{\boldsymbol{\rho}}_i(\epsilon) = \mathcal{L}[\hat{\boldsymbol{A}}](\hat{\boldsymbol{\rho}}_i(\epsilon))$, the Gaussian measurement acts as an effective Lindblad evolution that *decoheres* the system state with increasing ϵ via the action of $\hat{\boldsymbol{A}}$, but does not cause unitary disturbance along the natural flow of $\hat{\boldsymbol{A}}$ [83].

The exact expressions for the conditioned Gaussian detector means follow from (34) and (70),

$${}_f\langle x\rangle_T = g\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\{\hat{\boldsymbol{A}}, \hat{\boldsymbol{\rho}}_i\}))}{2\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))}, \tag{72a}$$

$${}_f\langle p\rangle_T = \frac{g}{i\hbar}\frac{\hbar^2}{4\sigma^2}\frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f(\mathrm{ad}\hat{\boldsymbol{A}})(\mathcal{E}(\hat{\boldsymbol{\rho}}_i)))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\mathcal{E}(\hat{\boldsymbol{\rho}}_i))}. \tag{72b}$$

Surprisingly, the special properties of the Gaussian moments (69) allow (72) to be written in a form proportional to the real and imaginary part of a complex weak-value involving the *decohered* system state (71) to all orders in the coupling strength g,

$$A_w(\epsilon) = \frac{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{A}}\hat{\boldsymbol{\rho}}_i(\epsilon))}{\mathrm{Tr}_s(\hat{\boldsymbol{P}}_f\hat{\boldsymbol{\rho}}_i(\epsilon))}, \tag{73a}$$

$${}_f\langle x\rangle_T = g\mathrm{Re}A_w(\epsilon), \tag{73b}$$

$${}_f\langle p\rangle_T = \frac{g}{\hbar}\frac{\hbar^2}{4\sigma^2}(2\mathrm{Im}A_w(\epsilon)). \tag{73c}$$

Following the interpretations outlined in this paper we can therefore understand the position shift $\mathrm{Re}A_w(\epsilon)$ to all orders in g as the *average of the observable* $\hat{\boldsymbol{A}}$ *in the decohered initial system state* $\hat{\boldsymbol{\rho}}_i(\epsilon)$ *conditioned on the post-selection* $\hat{\boldsymbol{P}}_f$. Similarly, we can understand the factor $2\mathrm{Im}A_w(\epsilon)$ in the momentum shift to all orders in g as the *logarithmic directional derivative of the probability of post-selecting* $\hat{\boldsymbol{P}}_f$ *given the decohered initial system state* $\hat{\boldsymbol{\rho}}_i(\epsilon)$ *along the unitary flow generated by* $\hat{\boldsymbol{A}}$.

If the measured operator is the qubit operator $\hat{\boldsymbol{A}} = A\hat{\boldsymbol{\sigma}}_3$ as in (53), then we can further simplify the expression (70) using the identities (55) to find,

$$\begin{aligned}\mathcal{E}(\hat{\boldsymbol{\rho}}_i) &= \hat{\boldsymbol{\rho}}_i + \frac{1}{2}(e^{-(Ag/\sigma)^2/2} - 1)\left(r_1\hat{\boldsymbol{\sigma}}_1 + r_2\hat{\boldsymbol{\sigma}}_2\right), \\ &= \frac{1}{2}\left(\hat{\mathbf{1}} + r_3\hat{\boldsymbol{\sigma}}_3 + e^{-(Ag/\sigma)^2/2}\left(r_1\hat{\boldsymbol{\sigma}}_1 + r_2\hat{\boldsymbol{\sigma}}_2\right)\right),\end{aligned} \tag{74}$$

which shows how the measurement decoheres the bases orthogonal to $\hat{\boldsymbol{A}}$ in the initial state with an increase in the dimensionless flow parameter $(Ag/\sigma)^2$ [84]. This decoherence is illustrated in Figure 4. The conditioned means

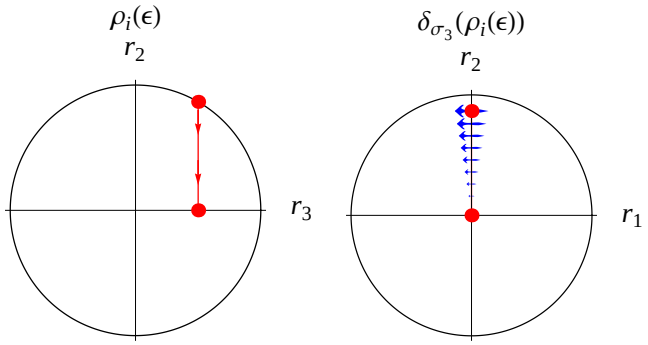

FIG. 4. (color online) Two projections of the Bloch sphere showing the pure decoherence of the specific state $\hat{\boldsymbol{\rho}}_i(\epsilon) = \exp(\epsilon\mathcal{L}[\hat{\boldsymbol{\sigma}}_3])(\hat{\boldsymbol{\rho}}_i) = (\hat{\mathbf{1}} + \exp(-2\epsilon)\sqrt{3}\hat{\boldsymbol{\sigma}}_2/2 + \hat{\boldsymbol{\sigma}}_3/2)/2$ due to the Gaussian detector (74). (left) The projection onto the plane $r_1 = 0$ showing the progressive collapse of $\hat{\boldsymbol{\rho}}_i(\epsilon)$ onto the r_3 axis with increasing ϵ. (right) The projection onto the plane $r_3 = 0$ showing the vector field $\delta_{\sigma_3}(\hat{\boldsymbol{\rho}}_i(\epsilon))$ during the progressive collapse. Notably the quantity $2\mathrm{Im}A_w(\epsilon)$ (73) is the rate of change of the post-selection probability (41) along this vector field for all ϵ, but not along the purely decohering trajectory that $\hat{\boldsymbol{\rho}}_i(\epsilon)$ actually follows.

(72) of a Gaussian qubit detector consequently have the exact form,

$$ {}_f\langle x\rangle_T = g\frac{\langle A\rangle_0}{\tilde{p}(f)}, \tag{75a}$$

$$ {}_f\langle p\rangle_T = \frac{g}{\hbar}\frac{\hbar^2}{4\sigma^2}\frac{2A}{\tilde{p}(f)}e^{-(Ag/\sigma)^2/2}\times \tag{75b}$$
$$ (r_1\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_2) - r_2\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_1)),$$

$$ \tilde{p}(f) = 1 + r_3\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_3) \tag{75c}$$
$$ + e^{-(Ag/\sigma)^2/2}\left(r_1\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_1) + r_2\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_2)\right),$$

to all orders in the coupling strength g. When expanded to linear order in g, (75) reduces to (38) with the real and imaginary parts of the qubit weak value (63), as expected.

For contrast, as g becomes large the unconditioned measurement of $\hat{\boldsymbol{A}}$ becomes essentially projective and the operation $\mathcal{E}$ almost completely decoheres the initial state (74) into the basis of $\hat{\boldsymbol{A}}$ as the pointer basis,

$$\mathcal{E}(\hat{\boldsymbol{\rho}}_i) \approx \frac{1}{2}\left(\hat{\mathbf{1}}_s + r_3\hat{\boldsymbol{\sigma}}_3\right). \tag{76}$$

Hence, in this *strong measurement* regime, the conditioned means (75) approximate,

$$ {}_f\langle x\rangle_T \approx g\frac{\langle A\rangle_0}{1 + r_3\mathrm{Tr}_s(\hat{\boldsymbol{\rho}}_f\hat{\boldsymbol{\sigma}}_3)}, \tag{77a}$$

$$ {}_f\langle p\rangle_T \approx 0. \tag{77b}$$

The position shift contains the average of $\hat{\boldsymbol{A}}$ in the *decohered* initial system state $\mathcal{E}(\hat{\boldsymbol{\rho}}_i)$, conditioned by the post-selection. Moreover, since the decohered initial system state $\mathcal{E}(\hat{\boldsymbol{\rho}}_i)$ is essentially diagonal in the basis of $\hat{\boldsymbol{A}}$, it will no longer Rabi oscillate, so the directional derivative along the flow generated by $\hat{\boldsymbol{A}}$ will be essentially zero. Hence, the probability correction factor represented by $2\mathrm{Im}A_w(\epsilon)$ vanishes.

VI. CONCLUSION

We have given an exact treatment of a conditioned von Neumann measurement for an arbitrary initial state and an arbitrary post-selection using the language of quantum operations. The full form of the conditioned detector response (34) naturally indicates how the measurement disturbance and conditioning from the post-selection modify the unconditioned detector response. The corresponding linear response of the detector (38) can be parametrized by the *generalized complex weak value* (35), but the origins of the real and imaginary parts differ.

The real part of the weak value (2) stems directly from the measurement of $\hat{\boldsymbol{A}}$ made on the initial system state, and can be interpreted as the idealized zero-disturbance conditioned average of the operator $\hat{\boldsymbol{A}}$ acting in its role as an observable. The imaginary part of the weak value (3), on the other hand, contains no information about the measurement of $\hat{\boldsymbol{A}}$ as an observable, but instead arises from the *disturbance* due to the von Neumann coupling. We interpret it as the logarithmic directional derivative (41) of the post-selection probability along the unitary flow in state space generated by the operator $\hat{\boldsymbol{A}}$ in its role as the element of a Lie algebra. The complex weak value therefore captures both halves of the dual role of the operator $\hat{\boldsymbol{A}}$ in the quantum formalism.

To illustrate this interpretation, we considered the weak value of momentum (51) post-selected on a particular position. Its real part is the Bohmian momentum representing the average momentum conditioned on a position detection, while its imaginary part (52) is proportional to the "osmotic velocity" that describes the logarithmic derivative of the probability density for measuring the particular position directed along the flow generated by the momentum.

Finally, we applied our exact solution to the useful special cases of a qubit operator and a Gaussian detector. We showed how the natural qubit Rabi oscillations that would be generated by the von Neumann interaction Hamiltonian (4) become disrupted by the measurement. We also showed that the Gaussian detector purely decoheres the initial system state into the basis of $\hat{\boldsymbol{A}}$ in the Lindblad sense, which allows the exact interpretation of the position and momentum shifts of the detector in terms of a complex weak value (73) that involves the *decohered* system state.

ACKNOWLEDGMENTS

This work was supported by the NSF Grant No. DMR-0844899, and ARO Grant No. W911NF-09-0-01417.

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988).

[2] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer, 1932).

[3] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, Phys. Rev. D **40**, 2112 (1989).

[4] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Phys. Rev. Lett. **66**, 1107 (1991).

[5] A. D. Parks, D. W. Cullin, and D. C. Stoudt, Proc. R. Soc. Lond. **454**, 2997 (1998).

[6] O. Hosten and P. Kwiat, Science **319**, 787 (2008).

[7] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Phys. Rev. Lett. **102**, 173601 (2009).

[8] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A **82**, 063822 (2010).

[9] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, Phys. Rev. A **82**, 011802(R) (2010).

[10] M. D. Turner, C. A. Hagedorn, S. Schlamminger, and J. H. Gundlach, Opt. Lett. **36**, 1479 (2011).

[11] J. M. Hogan, J. Hammer, S.-W. Chiow, S. Dickerson, D. M. S. Johnson, T. Kovachy, A. Sugarbaker, and M. A. Kasevich, Opt. Lett. **36**, 1698 (2011).

[12] R. Jozsa, Phys. Rev. A **76**, 044103 (2007).

[13] A. Di Lorenzo and J. C. Egues, Phys. Rev. A **77**, 042108 (2008).

[14] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A **80**, 041803 (2009).

[15] T. Geszti, Phys. Rev. A **81**, 044102 (2010).

[16] Y. Shikano and A. Hosoya, J. Phys. A **43**, 025304 (2010).

[17] Y.-W. Cho, H.-T. Lim, Y.-S. Ra, and Y.-H. Kim, New J. Phys. **12**, 023036 (2010).

[18] S. Wu and Y. Li, Phys. Rev. A **83**, 052106 (2011).

[19] E. Haapasalo, P. Lahti, and J. Schultz, Phys. Rev. A **84**, 052107 (2011).

[20] A. D. Parks and J. E. Gray, Phys. Rev. A **84**, 012116 (2011).

[21] X. Zhu, Y. Zhang, S. Pang, C. Qiao, Q. Liu, and S. Wu, Phys. Rev. A **84**, 052111 (2011).

[22] Y. Shikano and A. Hosoya, Physica E **43**, 776 (2011).

[23] A. G. Kofman, S. Ashkab, and F. Nori, "Nonperturbative theory of weak pre- and post-selected measurements," (2011), arXiv:1109.6315.

[24] K. Nakamura, A. Nishizawa, and M.-K. Fujimoto, "All-order evaluation of weak measurements: — The cases of an operator $\mathbf{A}$ which satisfies the property $\mathbf{A}^2 = 1$ —," (2011), arXiv:1108.2114.

[25] T. Koike and S. Tanaka, "On Amplification by Weak Measurement," (2011), arXiv:1108.2050.

[26] A. Di Lorenzo, "Full counting statistics of weak measurement," (2011), arXiv:1109.4183.

[27] A. K. Pan and A. Matzkin, "Weak and semiweak values in non-ideal measurements: an exact treatment," (2011), arXiv:1109.6824.

[28] A. Romito, Y. Gefen, and Y. M. Blanter, Phys. Rev. Lett. **100**, 056801 (2008).

[29] N. Brunner and C. Simon, Phys. Rev. Lett. **105**, 010405 (2010).

[30] O. Zilberberg, A. Romito, and Y. Gefen, Phys. Rev. Lett. **106**, 080405 (2011).

[31] S. Kagami, Y. Shikano, and K. Asahi, Physica E **43**, 761 (2011).

[32] C.-F. Li, X.-Y. Xu, J.-S. Tang, J.-S. Xu, and G.-C. Guo, Phys. Rev. A **83**, 044102 (2011).

[33] J. C. Howell, D. J. Starling, P. B. Dixon, P. K. Vudyasetu, and A. N. Jordan, Phys. Rev. A **81**, 033813 (2010).

[34] Y. Aharonov, F. Columbo, I. Sabadini, D. C. Struppa, and J. Tollaksen, J. Phys. A: Math. Theor. **44**, 365304 (2011).

[35] M. V. Berry and P. Shukla, J. Phys. A: Math. Theor. **45**, 015301 (2011).

[36] Y. Aharonov and L. Vaidman, Phys. Rev. A **41**, 11 (1990).

[37] Y. Aharonov and A. Botero, Phys. Rev. A **72**, 052111 (2005).

[38] J. Tollaksen, J. Phys. A **40**, 9033 (2007).

[39] A. Hosoya and M. Koga, J. Phys. A: Math. Theor. **44**, 415303 (2011).

[40] Y. Aharonov and L. Vaidman, Lect. Notes Phys. **734**, 399 (2008).

[41] A. M. Steinberg, Phys. Rev. Lett. **74**, 2405 (1995).

[42] A. M. Steinberg, Phys. Rev. A **52**, 32 (1995).

[43] Y. Aharonov, N. Erez, and B. Reznik, Phys. Rev. A **65**, 052124 (2002).

[44] Y. Aharonov, N. Erez, and B. Reznik, J. Mod. Opt. **50**, 1139 (2003).

[45] D. Rohrlich and Y. Aharonov, Phys. Rev. A **66**, 042102 (2002).

[46] H. M. Wiseman, Phys. Rev. A **65**, 032111 (2002).

[47] H. M. Wiseman, Phys. Lett. A **311**, 285 (2003).

[48] R. Mir, J. S. Lundeen, M. W. Mitchell, A. M. Steinberg, J. L. Garretson, and H. M. Wiseman, N. J. Phys. **9**, 287 (2007).

[49] N. Brunner, V. Scarani, M. Wegmuller, M. Legre, and N. Gisin, Phys. Rev. Lett. **93**, 203902 (2004).

[50] Y. Aharonov and L. Vaidman, J. Phys. A: Math. Gen. **24**, 2315 (1991).

[51] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, Phys. Lett. A **324**, 125 (2004).

[52] D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann, Phys. Rev. Lett. **92**, 043601 (2004).

[53] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, Phys. Lett. A **301**, 130 (2002).

[54] J. S. Lundeen and A. M. Steinberg, Phys. Rev. Lett. **102**, 020404 (2009).

[55] K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, New J. Phys. **11**, 033011 (2009).

[56] A. Hosoya and Y. Shikano, J. Phys. A **43**, 385307 (2010).

[57] Q. Wang, F. W. Sun, Y. S. Zhang, Jian-Li, Y. F. Huang, and G. C. Guo, Phys. Rev. A **73**, 023814 (2006).

[58] C. R. Leavens, Found. Phys. **35**, 469 (2005).

[59] H. M. Wiseman, New J. Phys. **9**, 165 (2007).

[60] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg, Science **332**, 1170 (2011).

[61] B. J. Hiley, "Weak Values: Approach through the Clifford and Moyal Algebras," (2011), arXiv:1111.6536.

[62] N. S. Williams and A. N. Jordan, Phys. Rev. Lett. **100**, 026804 (2008).

[63] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O'Brien, A. G. White, and G. J. Pryde, Proc. Natl. Acad. Sci. U. S. A. **108**, 1256 (2011).

[64] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N.

Jordan, Phys. Rev. Lett. **106**, 040402 (2011).
[65] G. J. Pryde, J. L. O'Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, Phys. Rev. Lett. **94**, 220405 (2005).
[66] M. Iinuma, Y. Suzuki, G. Taguchi, Y. Kadoya, and H. F. Hofmann, New J. Phys. **13**, 033041 (2011).
[67] J. Dressel, Y. Choi, and A. N. Jordan, "Measuring Which-Path Information with Coupled Electronic Mach-Zehnder Interferometers," (2011), arXiv:1105.2587.
[68] J. Dressel, S. Agarwal, and A. N. Jordan, Phys. Rev. Lett. **104**, 240401 (2010).
[69] J. Dressel and A. N. Jordan, J. Phys. A: Math. Theor. **45**, 015304 (2012).
[70] J. Dressel and A. N. Jordan, "Contextual Values Approach to the Generalized Measurement of Observables," (2011), arXiv:1110.0418.
[71] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[72] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[73] Note that such a value assignment does not violate the Bell-Kochen-Specker theorem [38, 71, 72] since (2) does not generally obey the product rule, $(AB)_w \neq A_w B_w$.
[74] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, 2000).
[75] H. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2007).
[76] H. M. Wiseman and G. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2009).
[77] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Phys. Rev **134**, B1410 (1964).
[78] D. Pegg, S. Barnett, and J. Jeffers, J. Mod. Opt. **49**, 913 (2002).
[79] T. Amri, J. Laurat, and C. Fabre, Phys. Rev. Lett. **106**, 020502 (2011).
[80] E. Nelson, Phys. Rev. **150**, 1079 (1966).
[81] D. Bohm and B. J. Hiley, Phys. Reports **172**, 93 (1989).
[82] G. Lindblad, Comm. Math. Phys. **48**, 119 (1976).
[83] That von Neumann coupling can lead to such Lindblad evolution in the case of *continuous* indirect non-selective measurements with zero-mean detectors was also noted by Breuer and Petruccione [75, §3.5.2, p.162]; however, here we make a single Gaussian detection after a duration of time, producing an effective flow parameter $(g/2\sigma)^2$.
[84] The continuous measurement of a double quantum dot that is discussed in [62, 68] can be mapped onto this problem so that $(Ag/\sigma)^2 \to \gamma t$ where $\gamma = 1/T_m$ is an inverse characteristic measurement time that acts as a dephasing rate due to the continuous non-selective measurement. Hence, the results obtained therein are special cases of the exact solution (75).