Weak values as interference phenomena

Justin Dressel
Department of Electrical and Computer Engineering, University of California, Riverside, California 92521, USA
(Received 3 October 2014; revised manuscript received 25 February 2015; published 18 March 2015)

Weak values arise experimentally as conditioned averages of weak (noisy) observable measurements that minimally disturb an initial quantum state, and also as dynamical variables for reduced quantum state evolution even in the absence of measurement. These averages can exceed the eigenvalue range of the observable ostensibly being estimated, which has prompted considerable debate regarding their interpretation. Classical conditioned averages of noisy signals only show such anomalies if the quantity being measured is also disturbed prior to conditioning. This fact has recently been rediscovered, along with the question whether anomalous weak values are merely classical disturbance effects. Here we carefully review the role of the weak value as both a conditioned observable estimation and a dynamical variable, and clarify why classical disturbance models will be insufficient to explain the weak value unless they can also simulate other quantum interference phenomena.

DOI: 10.1103/PhysRevA.91.032116 PACS number(s): 03.65.Ta, 03.67.−a, 42.25.Hz

I. INTRODUCTION

After their introduction over a quarter-century ago [1], quantum weak values [2,3] have consistently found themselves at the center of controversy [4,5]. Indeed, the original paper [1] details how one can postselect a weak (i.e., noisy) measurement of a spin-1/2 operator for an electron (using a sequence of two Stern-Gerlach apparatuses) to obtain a *conditioned expectation value* that approximates a weak value with an anomalously large value of 100. The question whether this strange average value has any physical meaning pertaining to the spin has since plagued the concept of the weak value (e.g., [6]).

The most recent addition to this controversy [7] considers a superficially similar example consisting of a classical coin that has its two faces noisily measured, then *disturbed*, and finally conditioned to produce an anomalous average value of 100 heads. The conclusion drawn from their study (which has been heavily criticized [8–12]) is that strange weak values may be understood entirely as classical disturbance effects, making them not "quantum." In fact, every element of this simple example of how intermediate disturbance can cause strange postselected averages of noisy signals has been previously demonstrated, and corroborates our published work: not only did we emphasize a similar disturbance example using a colored marble in our systematic investigation of generalized observable measurements [13,14], but we also carefully highlighted the potential role of invasive measurements in studies linking strange conditioned averages (including weak values) to violations of generalized Leggett-Garg inequalities [15–19] (which were designed to test for "quantum" behavior in macroscopic systems [20–22]). It is now well established that any hidden-variable model that can produce strange conditioned averages like the weak value must include some form of intermediate disturbance (see also [23,24]). The more interesting question to raise is not whether a particular strange conditioned average may be explained as classical disturbance, but rather whether such models of disturbance can also reproduce the complete behavior of the weak value as its physical parameters are varied.

In this paper, we revisit this question in order to dispel the abundant confusion about weak values still present in the literature, and emphasize that a strange weak value is nonclassical in precisely the same manner that a single quantum particle can be considered to be nonclassical. Specifically, strange weak values fundamentally arise from *interference* (i.e., superoscillations [25,26]), and thus also appear in any wavelike field theory, such as classical optics [27–29], in a straightforward way. In such a classical field theory, anomalous weak values do faithfully indicate physical wave properties, despite how counterintuitive their predictions may seem. For example, the orbital part of the Poynting vector field of optical vortex beams, or evanescent fields, can show anomalous local momentum distributions that are precisely equal to strange weak values [30–32]. Therefore, as with any quantum interference effect, only the fact that discrete and independent random events can be measured (as opposed to attenuated wave intensities) will distinguish whether the statistics producing a strange weak value are truly quantum mechanical in origin. (Alternatively, entangling the degrees of freedom of distinct particles will not have a simple classical field interpretation, e.g., [17].) Nevertheless, even in the case of discrete measurement events the large number of measurements needed to statistically resolve such a weak value still imply that it is best considered as a dynamical physical variable for the effective (classical) mean field, and not necessarily to each individual quantum particle *a priori* [32]. To emphasize these subtle points, we carefully review several complementary approaches to deriving and understanding the weak value, paying special attention to its role as an ideal estimate, an experimentally measurable conditioned average, and as a classical dynamical variable for reduced quantum state evolution. This detailed treatment aims to supplement the simplified introduction to the experimental applications of weak values in [2] with an expanded theoretical discussion that highlights their pervasive and underappreciated role throughout the quantum formalism.

In what follows we also emphasize the often overlooked connection between weak values and joint quasiprobability distributions (such as the Kirkwood-Dirac [33–38], Terletsky-Margenau-Hill [39–42], and the various Moyal phase-space distributions [43–46]) that determine conditioned observable estimates. Notably, to obtain a strange weak value outside the usual eigenvalue bounds, these joint quasiprobability distributions must become negative as a consequence of the

nonclassical quantum interference between probability amplitudes. The best known examples of this intrinsic negativity from quantum interference occur in the Moyal phase-space distributions, such as the Wigner distribution for quadratures, or the Glauber-Sudarshan P distribution for coherent state amplitudes, which show such negativity with nonclassical optical states [43–46]. Indeed, for single particles such negativity in quasiprobability distributions has been proven to be an equivalent notion of "nonclassicality" as the need for *contextual* hidden variables [47,48] (in the sense of Bell-Kochen-Specker [49,50]), and formally arises from the usual operator noncommutativity of quantum mechanics. This connection between contextuality and strange weak values was emphasized in a recent proof by Pusey [51], as well as an earlier study by Tollaksen [23], and is consistent with the established understanding that strange weak values arise fundamentally from (quantum) interference. It follows that if a classical model as in [7] could really mimic the detailed functional structure of the weak value, then it would also be able to simulate other features that are normally considered to be quantum mechanical. We further emphasize this latter point by reviewing the deep connections between weak values and the classical dynamical variables of the Hamilton-Jacobi formalism, both in its fully quantum generalization that is equivalent to the Schrödinger equation, and in the resulting classical limit.

This paper is organized as follows. In Sec. II we derive and discuss how the weak value is the best statistical estimate for the average (but unmeasured) observable value for the times between two known measurement events. In Sec. III we discuss three approaches for experimentally verifying the weak value as the appropriate such estimate: weak von Neumann coupling, weak generalized observable measurements, and as the physical dynamical variables for reduced state evolution. In Sec. IV we explicitly connect the weak value to quasiprobabilities, focusing on the Terletsky-Margenau-Hill, Kirkwood-Dirac, and Wigner distributions, and explicitly connect weak values to classical mean-field dynamical variables using the Hamilton-Jacobi quantum-classical correspondence. We conclude in Sec. V.

II. WEAK VALUES AS ESTIMATES

Most of the controversy surrounding weak values rests upon their common (but unnecessary) association with an alternative time-symmetric approach to the quantum theory that involves two state vectors [52–57]. In this time-symmetric approach, one forward-propagates a state vector $|i\rangle$ from an initial time 0 to t in the usual way; however, one also *back-propagates* a second state vector $\langle f|$ from a final time T to t. While the initial-state vector $|i\rangle$ corresponds to a preparation procedure, the final-state vector $\langle f|$ corresponds to a *postselection* procedure.

Interestingly, the best estimate [58–60] of the average (unmeasured) value for an observable $\hat{A}$ at any time t in the interval $[0,T]$ is then not the expectation value $A(t)=\langle i|\hat{U}_t^\dagger \hat{A}\hat{U}_t|i\rangle$ (which neglects the information about the postselection), but

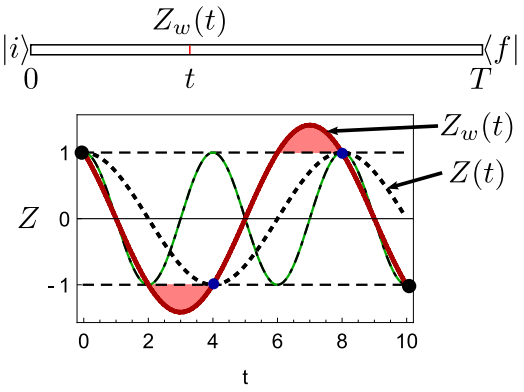

FIG. 1. (Color online) The best estimate of the average value of $\hat{Z}$ conditioned on two boundaries that bracket the time interval $[0,T]$ is the weak value $Z_w(t)=\mathrm{Re}\langle f|\hat{U}_{T-t}\hat{Z}\hat{U}_t|i\rangle/\langle f|\hat{U}_T|i\rangle$. For the Pauli $\hat{Z}=|1\rangle\langle 1|-|0\rangle\langle 0|$ qubit operator prepared in $|i\rangle=|1\rangle$, postselected in $\langle f|=\langle 0|$, and evolving with $\hat{U}_t=\exp[i\omega t(|1\rangle\langle 0|+|0\rangle\langle 1|)]$, the weak value $Z_w(t)$ (thin, green) coincides with the expectation value $Z(t)=\langle i|\hat{U}_t^\dagger\hat{Z}\hat{U}_t|i\rangle$ (dot-dashed, black) when the postselection is consistent with the natural oscillation. Otherwise, $Z(t)$ (dashed, black) displays a jump at time T, while $Z_w(t)$ (red) smoothly connects the boundaries, still passing through the same points of certain Z (blue dots). The shaded regions exceed the eigenvalue bounds of ± 1, indicating the inconsistency between the natural evolution and the observed boundaries.

is rather the *weak value*

$$A_w(t)=\mathrm{Re}\frac{\langle f|\hat{U}_{T-t}\hat{A}\hat{U}_t|i\rangle}{\langle f|\hat{U}_T|i\rangle}, \tag{1}$$

as we will derive shortly. Here $\hat{U}_t=\exp(-i\hat{H}t/\hbar)$ is the unitary propagator for a time interval t that is generated by the Hamiltonian $\hat{H}$. Note that the imaginary part of the weak value, while independently interesting as a measure of intrinsic measurement disturbance [59,61,62], is unrelated to the estimation of $\hat{A}$ as an observable, so we will ignore it for now.

The problematic feature of Eq. (1) as an estimate is that it may exceed the eigenvalue range of $\hat{A}$; such strange behavior is illustrated as the shaded areas in Fig. 1. As discussed in the Introduction, a classical conditioned estimate may show such anomalous behavior only if the estimation procedure is *noisy* and if what is being estimated is *disturbed* in the interval $[0,T]$ [14,15,17,23,24]. The question raised in Ref. [7] is whether such a classical model with noisy estimation and disturbance is sufficient to explain Eq. (1). As we will show in what follows, such a classical disturbance explanation is difficult to defend when the many roles of this weak value expression are examined in detail.

A. Derivation of the best estimate

For completeness we now show how this weak value formula can be obtained as the optimal estimate that minimizes a statistical uncertainty metric, justifying its interpretation as a best estimate.

Suppose one wishes to estimate the best average values for an observable $\hat{A}$ given a initial preparation $|i\rangle$, followed by a projective measurement in a particular basis $|f\rangle$ that

does not commute with $\hat{A}$. For each f, we can guess a value $\bar{a}_f$ that estimates the average of $\hat{A}$ given that the specific result f was observed. This procedure formally constructs a Hermitian observable $\hat{A}_{\text{est}} = \sum_f \bar{a}_f |f\rangle\langle f|$ that is measured by the procedure, which contains the estimates for every f. The goal is then to determine the optimal such estimates $\bar{a}_f^{\text{opt}}$ of $\hat{A}$ conditioned on each measured result f.

To accomplish this goal, we must define a measure for how close the estimate $\hat{A}_{\text{est}}$ is to the target observable $\hat{A}$. A natural choice for such a measure is the weighted trace distance between two observables [41,58,60]

$$\mathcal{D}_\rho(\hat{A},\hat{B}) = \text{Tr}(\hat{\rho}(\hat{A}-\hat{B})^2), \tag{2}$$

which can generally depend upon any positive prior bias $\hat{\rho}$. We can interpret this measure as specifying the shortest *geometric distance* in operator space, weighted by the *statistical prior* bias of $\hat{\rho}$. Since $\hat{A}$ and $\hat{B}$ need not commute, there is generally no straightforward operational interpretation for this distance [63–65], but it does formally provide a reasonable definition of the geometric "closeness" for the two operators.

Now suppose we have a definite prior state $\hat{\rho} = |i\rangle\langle i|$ and choose $\hat{B} = \hat{A}_{\text{est}}$. Computing the weighted trace distance yields [41,58,60]

$$\begin{aligned}\mathcal{D}_\psi(\hat{A},\hat{A}_{\text{est}}) &= \langle\psi|[\hat{A}^2 + \hat{A}_{\text{est}}^2 - (\hat{A}\hat{A}_{\text{est}} + \hat{A}_{\text{est}}\hat{A})]|\psi\rangle,\\ &= \langle i|\hat{A}^2|i\rangle - \sum_f |\langle f|i\rangle|^2\left[\text{Re}\frac{\langle f|\hat{A}|i\rangle}{\langle f|i\rangle}\right]^2\\ &\quad + \sum_f |\langle f|i\rangle|^2\left[\bar{a}_f - \text{Re}\frac{\langle f|\hat{A}|i\rangle}{\langle f|i\rangle}\right]^2.\end{aligned} \tag{3}$$

Only the final term depends on the choice of estimates $\bar{a}_f$ for each f, and is positive definite. Therefore, the trace distance is minimized when this term vanishes, which in turn implies that the optimal estimate for each independent f must be the weak value formula

$$\bar{a}_f^{\text{opt}} = \text{Re}\frac{\langle f|\hat{A}|i\rangle}{\langle f|i\rangle}. \tag{4}$$

Note that if the basis $|f\rangle$ is chosen to be the eigenbasis of $\hat{A}$, then these optimal estimates reduce identically to the eigenvalues of $\hat{A}$ and the trace distance vanishes.

To obtain Eq. (1) including intermediate time evolution, we can choose a particular $\langle f| \mapsto \langle f|\hat{U}_{T-t}$ and initial state $|i\rangle \mapsto \hat{U}_t|i\rangle$ that take into account Hamiltonian propagation $\hat{U}_t = \exp(t\hat{H}/i\hbar)$ from the observed results for a specific preparation and postselection.

This pure-state derivation may also be generalized in a straightforward way [60], which produces a formulation of the weak value suitable for generalized measurements and mixed states

$$A_w(t) = \text{Re}\frac{\text{Tr}(\hat{E}_{T-t}\hat{A}\hat{\rho}_t)}{\text{Tr}(\hat{E}_{T-t}\hat{\rho}_t)}. \tag{5}$$

Here the back-propagating operator $\hat{E}_{T-t}$ is often called a "retrodictive state," or "effect matrix," in contrast to the forward-propagating "predictive state" $\hat{\rho}_t$, or "density matrix." Recently, the estimate in Eq. (5) has been used to great effect experimentally [66–69] for "quantum smoothing" [70,71] and "past quantum state" analyses [72,73] of continuously measured signals (e.g., it was used to track individual photon emissions into a monitored cavity [69]). Both $\hat{E}_{T-t}$ and $\hat{\rho}_t$ generally evolve according to open-system master equations [74,75] that can also include the effects from additional (discrete or continuous-in-time) stochastic measurement results [72,76,77], in contrast to the closed-system (unitary) Schrödinger–von Neumann dynamics usually assumed with Eq. (1). Note that if the effect matrix $\hat{E}_{T-t}$ is the identity $\hat{1}$, then no posterior conditioning has been performed, so the usual expectation value is also recovered as a special case.

B. Interpreting and generalizing the estimate

As a philosophical side note, for those who believe that the state-vector represents the complete physical (ontic) reality (e.g., adherents to the many-worlds interpretation [78]), this time-symmetric estimate prompts several more radical speculations: the existence of the second state vector $\langle f|\hat{U}_{T-t}$ in Eq. (1) seems to imply not only that the state $\hat{U}_t|i\rangle$ is an incomplete description of reality at time t, but also that there seems to be a causal effect on the time t from the future time T [79]. Such a *retrocausal* interpretation is similar in spirit to the interpretations of antiparticles in quantum field theory as field excitations that move backwards through time [80]. However, just as with antiparticles, one does not need to invoke such controversial philosophical concepts as physical state vectors or retrocausation to meaningfully interpret the weak value in Eq. (1) as the best available statistical estimate given only the information about the specified boundary conditions.

A more pragmatic attitude (which we shall adopt here) is to treat the estimate in Eq. (1) as *subjective* (epistemic), and pertaining to a time interval $[0,T]$ that has already occurred in the past. That is, one performs an experiment that prepares $|i\rangle$ at time 0, waits a duration T, then makes a projective measurement that shows a result corresponding to the state $\langle f|$. One then interprets Eq. (1) as the best estimate of the (unmeasured) average value of $\hat{A}$ within that time interval [58–60], given only the knowledge of $|i\rangle$, $\langle f|$, and $\hat{H}$. We emphasize that this approach is no different in character than stating that the expectation value $\langle i|\hat{U}_t^\dagger\hat{A}\hat{U}_t|i\rangle$ is the best estimate for the (unmeasured) average value of $\hat{A}$, given only the knowledge of the preparation $|i\rangle$ and $\hat{H}$. Indeed, such a counterfactual interpretation of the expectation value in the absence of measurement is at the core of the Ehrenfest theorem that equates quantum expectation values with mean-field classical dynamical variables [81]. If we interpret weak values as similar (but additionally constrained) classical dynamical variables, then an anomalous weak value should indicate the presence of some interesting intermediate physical process that must have occurred in order to satisfy both boundary conditions that bracket the time interval $[0,T]$ (see Fig. 1).

Supporting this point of view is the fact that similar bidirectional (in time) estimates about unknown properties of structured stochastic processes (e.g., hidden Markov models) during such an interval are now well established in classical computational mechanics [82–85]. There it is shown that one should use both forward and reverse "causal states" (i.e., probability distributions) that contain information gathered

both before *and after* each time t to optimally estimate the properties of an evolving stochastic process. Similarly, classical statistics and filtering theory also use bidirectional states to provide the best estimate for information contained in noisy data confined to a time interval (called optimally "smoothing" the noise) [86,87]. Since quantum theory is closely related to probability theory [14,88], it is logical that similar estimation methods can be applied. Indeed, upgrading these estimation schemes to the quantum realm [76,88,89] produces both states in Eq. (1), as well as the mixed-state generalization of Eq. (5) [90–92].

III. MEASURING WEAK VALUES

The confidence that estimations like the expectation value, or the weak value in Eqs. (1) and (5), reflect something meaningful about the physical world (and are not merely fevered hallucinations of the mind) follows from verification of their predictions by experimental measurements. In the case of the expectation value, any unbiased estimation of $\hat{A}$ will suffice, corroborating the predicted result. In the case of the weak value, however, the presence of the posterior boundary condition additionally constrains the form of the possible measurements that can verify the estimate.

Specifically, those measurements must be "weak," meaning that they should not appreciably perturb the evolution of the quantum system. Since information extraction necessarily disturbs the quantum state, only minimally informative (i.e., *noisy*) measurements will leave the state mostly unperturbed [1,75], and thus faithfully reproduce the assumptions made about the evolution during the time interval $[0,T]$ by the formulas in Eqs. (1) and (5). The surprising fact is that averaging such weak observable measurements can indeed consistently verify the weak value as the correctly estimated average, even when it predicts anomalous averages.

In what follows we will detail the standard von Neumann approach for measuring the weak value, as well as a more general approach that solidifies its interpretation as a conditioned average in the limit of negligible disturbance to the quantum state. We will also detail how the weak value appears as a classical dynamical variable for reduced system evolution even outside the usual context of postselected weak measurements, further cementing its interpretation as the appropriate mean-field variable that physically estimates the observable when the natural state evolution is unchanged by external influences.

A. von Neumann interaction

The standard approach for performing a weak observable measurement [1–3,93], originally due to von Neumann [94], is to couple the *system* observable of interest $\hat{A}$ (such as the spin of a particle) to a *detector* observable $\hat{F}$ (such as the transverse momentum of the same particle) for an independent degree of freedom, using a simple linear interaction Hamiltonian

$$\hat{H}_{DS}(t) = \hbar g(t) \hat{F} \otimes \hat{A}. \tag{6}$$

The time-dependent coupling profile $g(t)$ is typically assumed to be zero outside a short interval of duration δt and to be impulsive, i.e., short on the time scale of the natural dynamics of both the system and the detector. In the interaction picture for the system and detector, this coupling Hamiltonian produces a joint unitary rotation of the joint state that entangles the system with the detector

$$\hat{V}_{DS} = \exp(-i g \hat{F} \otimes \hat{A}), \tag{7}$$

where $g = \int_0^{\delta t} g(t')dt'$ is the effective coupling strength for the impulsive interaction.

For a concrete example of such an interaction, in the optical experiments [27–29] a wafer of birefringent crystal was used to couple the polarization (system) of a paraxial beam to its continuous transverse momentum (detector). In contrast, the optical experiments [16,17,95] used polarization-dependent reflection from partially transmitting optical elements to couple the polarization (system) of a paraxial beam to its binary orbital which-path degree of freedom (detector). In the more recent superconducting qubit experiment [18] the energy basis of one transmon qubit (system) was coupled via an intermediate bus stripline resonator to the binary energy basis of a second and physically separated transmon qubit (detector) using microwave pulses.

To derive the effect of such an interaction, consider the coupling procedure in Fig. 2. Suppose that the impulsive coupling begins at a time t, and that the joint state of the detector-system degrees of freedom is initially a product state $|d'\rangle|i'\rangle$, where the initial detector and system states at time t

$$|d'\rangle = \hat{U}^{(D)}_{t-t_0}|d\rangle, \quad |i'\rangle = \hat{U}_t|i\rangle \tag{8}$$

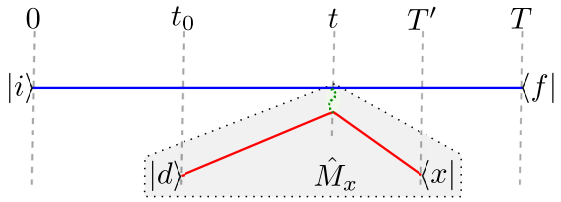

FIG. 2. (Color online) Impulsive von Neumann interaction for measuring $\hat{A} = \sum_a a|a\rangle\langle a|$. A system degree of freedom $|i\rangle$ (blue, top) is prepared at time 0, and a detector degree of freedom $|d\rangle$ (red, bottom) is prepared at a potentially different time t_0. After independent unitary propagation $\hat{U}_t$ and $\hat{U}^{(D)}_{t-t_0}$, respectively, to the intermediate time t they interact impulsively (green, wavy) with a joint unitary interaction $\hat{V}_{DS}$ for a short duration δt. After the interaction, both system and detector continue to propagate until the detector is measured at a time T' to find the result $\langle x|$, and the system is measured at a potentially different time T to find the result $\langle f|$. This entire coupling procedure may be represented as a single system operation (lightly shaded region) $\hat{M}_x = \langle x|\hat{U}^{(D)}_{T'-t}\hat{V}_{DS}\hat{U}^{(D)}_{t-t_0}|d\rangle$ that includes the detector preparation, evolution, coupling, and measurement together, producing a net effect only on the evolving system state. If the identity $\hat{A} = \sum_x \alpha_x \hat{M}_x^\dagger \hat{M}_x$ can be satisfied for some calibrated signal values α_x for the detector, then the system observable $\hat{A}$ can be faithfully estimated by the detector results x in an unbiased way. For strong measurements by the detector, the effective system operations become eigenstate projections $\hat{M}_x \to |a\rangle\langle a|$ with an eigenvalued signal $\alpha_x \to a$, while for weak measurements the operations approximate the identity $\hat{M}_x \approx \hat{1}$ with a noisy signal $|\alpha_x| \gg |a|$, leaving the state essentially unperturbed. Averaging this noisy signal α_x for such weak measurements conditioned on a particular f approximates the weak value $\mathrm{Re}\langle f|\hat{U}_{T-t}\hat{A}\hat{U}_t|i\rangle/\langle f|\hat{U}_T|i\rangle$.

may have propagated from previous states $|i\rangle$ and $|d\rangle$ that were prepared at the possibly different earlier times $t = 0$ and $t = t_0$, respectively, following the independent Hamiltonian evolution

$$\hat{U}_t^{(D)} = \exp(-it\hat{H}_D/\hbar), \quad \hat{U}_t = \exp(-it\hat{H}_S/\hbar). \tag{9}$$

After the entangling interaction of Eq. (7), the detector and system are allowed to again evolve freely for (potentially different) times $T - t$ and $T' - t$, respectively, after which they are independently measured projectively in the bases $\langle x|$ and $\langle f|$ to obtain a pair of detector-system results (x, f). Back-propagating these measured states from T' and T to the time immediately following the impulsive coupling $t + \delta t \approx t$ produces the effective final states

$$\langle x'| = \langle x|\hat{U}_{T'-t}^{(D)}, \quad \langle f'| = \langle f|\hat{U}_{T-t}. \tag{10}$$

The joint probability distribution for the results (x, f) can then be written in the compact (scattering) form

$$p_{x,f} = |\langle x', f'|\hat{V}_{DS}|d', i'\rangle|^2. \tag{11}$$

Importantly, changing the durations of time T and T' before measuring the system and detector does not affect the general form of the joint distribution in Eq. (11); only the back-propagated states $|x'\rangle$ and $|f'\rangle$ will change from the free evolution when these durations are varied. For a concrete example of this effect, in the optical case of [27–29] the free evolution of the transverse momentum (detector) produces diffraction effects after the birefringent crystal, which alters the effective back-propagated state implied by a later transverse position measurement.

For a sufficiently small coupling strength g, the interaction only *weakly* perturbs the initial system states, and we can expand the joint interaction $\hat{V}_{DS}$ perturbatively. To good approximation, we find that the relative change of the joint probability in Eq. (11) due to the weak interaction has the form

$$\frac{p_{x,f}}{|\langle x', f'|d', i'\rangle|^2} \approx 1 + 2g\,\mathrm{Im}F_w A_w + g^2|F_w|^2|A_w|^2, \tag{12}$$

which involves only the *first-order* (complex) detector and system weak values

$$F_w = \frac{\langle x'|\hat{F}|d'\rangle}{\langle x'|d'\rangle}, \quad A_w = \frac{\langle f'|\hat{A}|i'\rangle}{\langle f'|i'\rangle}. \tag{13}$$

Note that continuing this expansion will produce an infinite series characterized entirely by higher-order weak values involving all powers of $\hat{A}$ and $\hat{F}$ [2]; however, this truncation that involves only the first-order weak values is remarkably accurate for sufficiently small coupling strengths that satisfy $g|F_w||A_w| \ll 1$ [96].

This joint distribution also determines the relative change of the marginalized distributions for the detector, $p_x = \sum_f p_{x,f}$, and system, $p_f = \sum_x p_{x,f}$, statistics alone

$$\frac{p_x}{|\langle x'|d'\rangle|^2} \approx 1 + 2g\langle A\rangle\mathrm{Im}F_w + g^2\langle A^2\rangle|F_w|^2, \tag{14}$$

$$\frac{p_f}{|\langle f'|i'\rangle|^2} \approx 1 + 2g\langle F\rangle\mathrm{Im}A_w + g^2\langle F^2\rangle|A_w|^2. \tag{15}$$

Critically, note that the reduced *detector* statistics involve the expectation value of the *system* operator $\langle A\rangle = \langle i'|\hat{A}|i'\rangle$, as well as its second moment $\langle A^2\rangle = \langle i'|\hat{A}^2|i'\rangle$. This dependence means that by examining only the statistics of the detector, we can *indirectly estimate* the system observable $\hat{A}$. Similarly, the system statistics involve the expectation value of the detector operator $\langle F\rangle = \langle d'|\hat{F}|d'\rangle$, as well as its second moment $\langle F^2\rangle = \langle d'|\hat{F}^2|d'\rangle$, so can be used to indirectly estimate the detector observable $\hat{F}$ in a symmetric way.

To perform such an estimation using only the detector statistics in the linear response regime (where we can neglect the terms that are second order in g), an experimenter weights each of the outcomes x of the detector by some scaling value α_x, which constructs an effective detector readout observable [3,13]

$$\hat{R} = \sum_x \alpha_x |x\rangle\langle x|, \tag{16}$$

and produces the detector average (keeping all system results)

$$\sum_x \alpha_x p_x \approx \langle R\rangle + 2g\langle A\rangle\mathrm{Im}\langle RF\rangle, \tag{17}$$

in terms of the detector quantities $\langle R\rangle = \langle d'|\hat{R}|d'\rangle$ and $\langle RF\rangle = \langle d'|\hat{R}\hat{F}|d'\rangle$. Hence, by calibrating a known initial detector state $|d'\rangle$ and choosing the observables $\hat{F}$ and $\hat{R}$ strategically, one can extract the expectation value $\langle A\rangle$ in an unbiased way using only the detector statistics. Note that typically the detector state $|d'\rangle$ is chosen such that the relevant observables have zero mean prior to the interaction, $\langle F\rangle = \langle R\rangle = 0$.

Now suppose our experimental setup additionally filters the system outcomes so that only f may occur. (Alternatively, we can select only those particular events in the postprocessing of data that includes more outcomes.) The statistics of the laboratory detector x measurements that are properly conditioned on a particular system f outcome will then have the usual form from Bayes' theorem

$$p_{x|f} = \frac{p_{x,f}}{\sum_x p_{x,f}}. \tag{18}$$

The approximate relative change of this conditional distribution is thus

$$\frac{p_{x|f}}{|\langle x'|d'\rangle|^2} \approx \frac{1 + 2g\,\mathrm{Im}F_w A_w + g^2|F_w|^2|A_w|^2}{1 + 2g\langle F\rangle\mathrm{Im}A_w + g^2\langle F^2\rangle|A_w|^2}. \tag{19}$$

When the terms of order g^2 can be neglected, and when $\langle F\rangle = 0$, we thus have the linear response result that can be compared to the unfiltered estimation in Eq. (17)

$$\begin{aligned}\sum_x \alpha_x p_{x|f} &\approx \langle R\rangle + 2g\,\mathrm{Im}A_w\langle RF\rangle, \\ &= \langle R\rangle + 2g[\mathrm{Re}A_w\mathrm{Im}\langle RF\rangle + \mathrm{Im}A_w\mathrm{Re}\langle RF\rangle].\end{aligned} \tag{20}$$

Importantly, the weak value factor $\mathrm{Re}A_w$ that scales $\mathrm{Im}\langle RF\rangle$ corresponds directly to the expectation value $\langle A\rangle$ in Eq. (17). That is, the filtering procedure partitions the total average $\langle A\rangle$ into subensembles that have conditioned averages of $\mathrm{Re}A_w$, which precisely matches what we expect from the best estimate of $\hat{A}$ in Eq. (1). The final term involving $\mathrm{Im}A_w$ averages to zero in the total ensemble, and corresponds to the intrinsic symmetric backaction of the detector on the system due to the joint interaction of Eq. (7); it does not correspond to the estimation of $\hat{A}$ [61,97,98], and can be removed from the detector signal

in practice while preserving the estimation of $\langle A\rangle$ by choosing the detector observables such that $\mathrm{Re}\langle RF\rangle = 0$.

B. Observable estimation

The preceding discussion of the von Neumann coupling is traditional in the weak value literature [1–3] and is often sufficient for describing experimental implementations that measure weak values (e.g., [27–29]). However, this perturbative derivation makes the association between the estimations of the total observable average $\langle A\rangle$ and the real part of the weak value $\mathrm{Re}\,A_w$ somewhat inferential, prompting skepticism about the appropriateness of the connection. Since $\mathrm{Re}\,A_w$ has such counterintuitive properties, having a more direct link between the estimation of $\hat{A}$ and $\mathrm{Re}\,A_w$ as a conditioned average value more generally is desirable. Thankfully, we can indeed demonstrate that this is the proper association by rephrasing the conditional estimation procedure using the formalism of generalized measurements [13,14,76].

To do this, we rewrite the joint probability of Eq. (11) entirely in the system space

$$p_{x,f} = |\langle f'|\hat{M}_x|i'\rangle|^2, \tag{21}$$

by defining a *measurement (Kraus) operator* that encodes the entire coupling and measurement procedure into a single system operator [75,99]

$$\hat{M}_x = \langle x'|\hat{V}_{DS}|d'\rangle. \tag{22}$$

Notice that this measurement operator has the form of a partial matrix element, but only contracts out the detector part of the joint unitary $\hat{V}_{DS}$ to leave a purely system operator, which directly corresponds to how the complete measurement *procedure* affects the system. Intuitively, for strong measurements of $\hat{A}$, the measurement operators $\hat{M}_x \to |a\rangle\langle a|$ will become projectors onto the eigenstates of $\hat{A}$, while for *weak* measurements (in the sense of [1]) the measurement operators $\hat{M}_x \approx \hat{1}$ will approximate the identity operator for all x, which leaves the initial state nearly unperturbed.

It follows that the marginalized distribution of only the detector results can be written

$$p_x = \sum_f p_{x,f} = \langle i'|\hat{M}_x^\dagger \hat{M}_x|i'\rangle, \tag{23}$$

which has the form of the *system* expectation of a probability operator

$$\hat{P}_x = \hat{M}_x^\dagger \hat{M}_x \tag{24}$$

for the *detector* result x. These probability operators are positive, and form a resolution of the identity in the system space

$$\sum_x \hat{P}_x = \langle d'|\hat{V}_{DS}^\dagger \left[\sum_x |x'\rangle\langle x'|\right] \hat{V}_{DS}|d'\rangle = \hat{1}, \tag{25}$$

making them a *probability operator-valued measure* (POM, or POVM). Such a POM is the operator version of a properly normalized probability distribution. Indeed, when $\hat{P}_x$ commutes with $\hat{A}$, its diagonal elements will be precisely the classical conditional probabilities $p_{x|a}$ describing the likelihoods of each detector result x given a definite preparation of a [14]. For each a, these probabilities then independently satisfy $\sum_x p_{x|a} = 1$ according to Eq. (25).

Generally speaking, any purity-preserving generalized measurement can be expressed as such a measurement operator $\hat{M}_x$ and associated POM $\hat{P}_x = \hat{M}_x^\dagger \hat{M}_x$ [75,99], with the von Neumann interaction being a special case for implementing such a measurement with a concrete Hamiltonian. For example, the polarization-dependent reflection measurement in [17] did not use a von Neumann Hamiltonian description of the interaction of the entangled photon pair with the glass microscope coverslip in the experimental analysis, but rather characterized the relevant POM operators $\hat{P}_x$ directly with a series of separate calibration measurements by *measuring* the conditional probabilities $p_{x|a}$ for known preparations of the eigenstates $|a\rangle$. Up to additional phases that were also calibrated in separate measurements, the diagonal elements of $\hat{M}_x$ then had the form $\sqrt{p_{x|a}}$. The benefit of this direct approach is that the effect of the actual measurement procedure may be experimentally measured, without requiring a more detailed model of the extended detector space.

Now suppose that we can use the measured probabilities p_x to estimate the expectation value of $\hat{A}$ in an unbiased way. To do this, we must weight the outcomes x of the detector with appropriately scaled values α_x (e.g., by rescaling a low-visibility signal in the usual way) [13]

$$\sum_x \alpha_x p_x = \langle i'|\left[\sum_x \alpha_x \hat{P}_x\right]|i'\rangle. \tag{26}$$

Evidently, to produce the expectation value $\langle A\rangle = \langle i'|\hat{A}|i'\rangle$ for any initial state $|i'\rangle$, we must be able to choose appropriate values α_x that will calibrate the detector to probe the generalized spectral expansion [13,14]

$$\hat{A} = \sum_x \alpha_x \hat{P}_x. \tag{27}$$

If it can be arranged, this operator identity will guarantee that the values α_x and measurement operators $\hat{M}_x$ (and thus the probability operators $\hat{P}_x$) will directly estimate $\hat{A}$ in an unbiased way. As a concrete example, in [1] the detector was chosen such that $\hat{F} = \hat{p}$ was the momentum operator, $\hat{R} = \hat{x}$ was the position operator, and $\langle x|d'\rangle = (2\pi\sigma^2)^{-1/4}\exp(-x^2/4\sigma^2)$ was a zero-mean Gaussian distribution, such that $\langle R\rangle = \langle F\rangle = 0$, $F_w = ix/2\sigma^2$, and thus $\langle \hat{R}\hat{F}\rangle = i/2$ in Eq. (17); setting the simple scaled values of $\alpha_x = x/g$ then directly yields an unbiased estimation of $\langle A\rangle$ for all g. Note, however, that this common choice for Gaussian measurements is generally not a unique choice for producing an unbiased estimation, since the dimension of the detector often exceeds that of the system [13,14].

Once we fix the weights α_x to achieve the estimation identity of Eq. (27), we also fix the partial average for each postselection f according to Eq. (21)

$$\sum_x \alpha_x p_{x,f} = \langle i'|\hat{O}_t|i'\rangle, \tag{28}$$

which we write compactly in terms of the operator

$$\begin{aligned}\hat{O}_t &\equiv \sum_x \alpha_x \hat{M}_x^\dagger |f'\rangle\langle f'|\hat{M}_x, \\ &= \frac{1}{2}(\hat{A}|f'\rangle\langle f'| + |f'\rangle\langle f'|\hat{A}) + \sum_x \alpha_x \mathcal{L}[\hat{M}_x^\dagger]|f'\rangle\langle f'|,\end{aligned} \tag{29}$$

that we write as a symmetric (Jordan) product [100] between the operator $\hat{A}$ and the postselection $|f'\rangle\langle f'|$, modified by a sum of weighted Lindblad (dissipation) operations [76]

$$\mathcal{L}[\hat{M}_x^\dagger](\cdot) \equiv \tfrac{1}{2}([\hat{M}_x^\dagger,\cdot]\hat{M}_x + \hat{M}_x^\dagger[\cdot,\hat{M}_x]), \tag{30}$$

familiar from open-system dynamics [74,75]. These Lindblad terms quantify the perturbation introduced by the measurement. Note that for *weak* measurements (i.e., $\hat{M}_x \approx \hat{1}$) these terms approximately vanish to leave only the symmetric product, signifying that both the initial system state and postselection are essentially unaffected by the measurement of $\hat{M}_x$ [14,76].

Expanding the partial average in Eq. (28) produces

$$\sum_x \alpha_x p_{x,f} = \mathrm{Re}\langle f'|\hat{A}|i'\rangle\langle i'|f'\rangle + \mathcal{E}[\alpha], \tag{31}$$

where we have introduced the Lindblad error terms

$$\mathcal{E}[\alpha] = \sum_x \alpha_x \langle i'|(\mathcal{L}[\hat{M}_x]|f'\rangle\langle f'|)|i'\rangle \tag{32}$$

that are produced entirely by the perturbation from the measurement. Conditioning this partial average on obtaining a particular f then yields

$$\frac{\sum_x \alpha_x p_{x,f}}{\sum_x p_{x,f}} = \frac{\mathrm{Re}\langle f'|\hat{A}|i'\rangle\langle i'|f'\rangle + \mathcal{E}[\alpha]}{\langle f'|i'\rangle\langle i'|f'\rangle + \mathcal{E}[1]}. \tag{33}$$

When the error terms $\mathcal{E}$ are small enough to be neglected [14,101] (meaning that the initial system state is negligibly perturbed), the *real part* of the weak value in Eq. (1) is unambiguously recovered as the *measured* conditioned estimate for $\hat{A}$, verifying our derivation of this real part as a best estimate. As expected, the imaginary part is unrelated to the estimation of $\hat{A}$, so does not contribute to Eq. (33), which justifies our interpretation of the terms in the von Neumann linear response of Eq. (20). Deriving Eq. (5) as a measured estimation is a similar exercise [93].

Importantly, nothing about the derivation of Eq. (33) changes when the time t, the initial system state $|i\rangle$, the system postselection $\langle f|$, or even the system Hamiltonian $\hat{H}$ are varied, as long as one keeps fixed the measurement procedure set by the choice of calibration weights α_x and corresponding $\hat{M}_x$ (e.g., the detector states $|d'\rangle$ and $\langle x'|$, Hamiltonian $\hat{H}^{(D)}$, and coupling interaction $\hat{V}_{DS}$). This robustness of the derivation implies that the same weak measurement procedure can approximate the *entire functional dependence* of the weak value in Eq. (1), in contrast to the single arbitrary value produced by the coin disturbance scheme in Ref. [7]. Moreover, the weak value in Eq. (1) no longer depends upon the specific measurement procedure, just like the expectation value in Eq. (26), so this approximation will work for *any* unbiased weak measurement procedure. The only requirement for consistently recovering the weak value Eq. (1) as the limiting value of the conditioned average in Eq. (33) is for the Lindblad perturbation terms in Eq. (30) to be small enough to neglect [76,93], meaning that the quantum state is approximately unperturbed. (For a physical example where the state disturbance from the coupling may not always be neglected, see [102].)

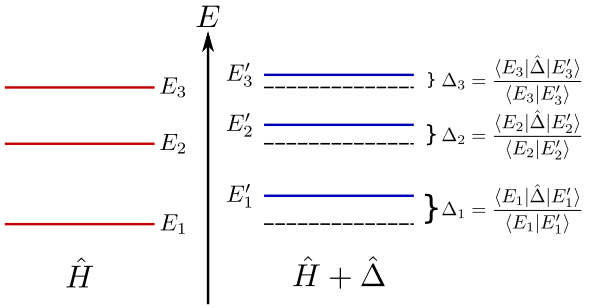

FIG. 3. (Color online) Energy-level perturbations. The eigenvalues E of a Hamiltonian $\hat{H}$ are shifted to new eigenvalues E' when a Hamiltonian perturbation $\hat{\Delta}$ is added. These shifts in energy levels are exactly (real) weak values $\langle E|\hat{\Delta}|E'\rangle/\langle E|E'\rangle$, and thus may lie outside the spectrum of the perturbation $\hat{\Delta}$.

C. Dynamical weak values

Thus far we have carefully discussed the most well-known role of weak values in experiment, namely as complex parameters that characterize a von Neumann interaction, and as conditioned estimations of observable averages. However, these are not the only places that weak values naturally appear. Here we consider two more common cases that are usually overlooked: weak values as eigenvalue perturbations, and weak values as classical dynamical variables in reduced system dynamics.

The first case is mathematically trivial to show, but has nontrivial implications. Consider the case of a Hamiltonian $\hat{H}$ with energy eigenstates $|E\rangle$ and eigenvalues $\hat{H}|E\rangle = E|E\rangle$. Suppose this Hamiltonian becomes perturbed by a new contribution $\hat{\Delta}$, producing new eigenstates $(\hat{H}+\hat{\Delta})|E'\rangle = E'|E'\rangle$. This latter eigenvalue equation can then be contracted with an unperturbed eigenstate $\langle E|$ and rearranged to find the following relation between the eigenvalues:

$$E' = E + \frac{\langle E|\hat{\Delta}|E'\rangle}{\langle E|E'\rangle}, \tag{34}$$

which is illustrated in Fig. 3. That is, the (purely real) weak value of the perturbation $\hat{\Delta}$ determines the shift in energy for the eigenstates of the Hamiltonian, and thus may lie outside the spectrum of $\hat{\Delta}$. One can understand this weak value as the best estimate of the average energy perturbation required to move from the old eigenstate $|E\rangle$ to the new eigenstate $|E'\rangle$. No measurement is being performed here, so this shift constitutes a dynamical effect where the weak value indeed represents the physical energy shift. This shift can be verified by measuring the eigenenergies before and after such a perturbation is added to the system.

The second case of a dynamical weak value is best illustrated by an explicit example, shown in Fig. 4. Suppose we couple a qubit dispersively to a single-mode resonator (e.g., a superconducting qubit setup like the one used in [18]). The simplest Hamiltonian for how such a joint system naturally evolves is

$$\hat{H} = \frac{\hbar\omega_q}{2}\hat{\sigma}_z + \hbar\omega_r\hat{a}^\dagger\hat{a} + \hbar\chi\hat{\sigma}_z\hat{a}^\dagger\hat{a}, \tag{35}$$

where $\hat{\sigma}_z = |1\rangle\langle 1| - |0\rangle\langle 0|$ is the Pauli Z operator between the qubit energy levels, $\hat{a}$ is the lowering operator of the resonator

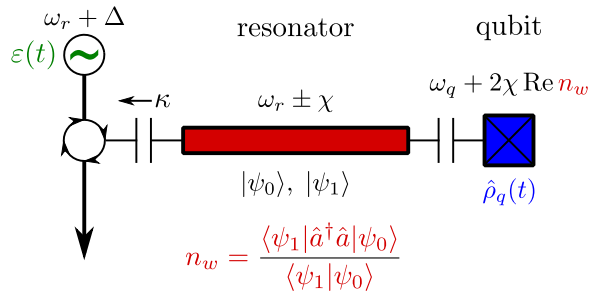

FIG. 4. (Color online) Superconducting qubit capacitively coupled to a one-sided stripline resonator with energy-decay rate κ, which is pumped through a circulator with a coherent microwave source $\varepsilon(t)$ at a frequency detuned by Δ from the bare resonator frequency ω_r. The resonator frequency is shifted by $\pm\chi$ depending on the qubit state due to the dispersive coupling. As such, each definite qubit state $|0\rangle$ or $|1\rangle$ approximately correlates to a distinct resonator state $|\psi_0\rangle$ or $|\psi_1\rangle$. The reduced qubit state $\hat{\rho}_q(t)$ then displays coherence oscillations at a frequency $\omega_q + 2\chi\,\mathrm{Re}n_w(t)$ due to the ac Stark shift, which depends on the real part of the weak value $n_w(t) = \langle\psi_1|\hat{a}^\dagger\hat{a}|\psi_0\rangle/\langle\psi_1|\psi_0\rangle$ of the resonator population. The qubit coherence similarly displays decay at an average rate $\Gamma = 2\chi\,\mathrm{Im}n_w(t)$ that depends on the imaginary part of $n_w(t)$, indicating measurement dephasing from ensemble averaging the fluctuations of the resonator population. Importantly, the complex weak value $n_w(t)$ is *physically* the relevant mean-field classical dynamical variable for the resonator population that affects the reduced qubit state at every point in time t, with the real part corresponding to the best classical estimation of the ensemble-averaged resonator population probed by coherent qubit superpositions of $|0\rangle$ and $|1\rangle$.

mode satisfying $[\hat{a},\hat{a}^\dagger] = 1$, ω_q and ω_r are the oscillation frequencies of the qubit and resonator, and $\pm\chi$ is the dispersive frequency shift of the resonator that depends on the qubit state. Note that the interaction term between the qubit and the resonator has the general von Neumann form of Eq. (6), but it is no longer impulsive.

Assuming a pure state for the joint qubit-resonator system, we can make the following ansatz for the form of the joint state

$$|\Psi\rangle = c_0(t)\,|0\rangle|\psi_0(t)\rangle + c_1(t)\,|1\rangle|\psi_1(t)\rangle, \tag{36}$$

where $c_{0,1}(t)$ are complex amplitudes, and $|\psi_{0,1}(t)\rangle$ are normalized resonator states that are correlated to each definite qubit state. It follows that the reduced qubit density matrix $\hat{\rho}_q = \mathrm{Tr}_r|\Psi\rangle\langle\Psi|$ after tracing out the resonator has the following diagonal populations

$$p_0(t) = |c_0(t)|^2, \quad p_1(t) = |c_1(t)|^2, \tag{37}$$

as well as an off-diagonal coherence

$$\rho_{01}(t) = c_1^*(t)c_0(t)\langle\psi_1(t)|\psi_0(t)\rangle, \tag{38}$$

that depends explicitly on the overlap between the two distinct and dynamically evolving resonator states that are correlated to definite qubit populations.

The simple Hamiltonian considered in Eq. (35) is already diagonal in the energy basis of the qubit, so the populations in Eq. (37) do not change in time. However, the coherence in Eq. (38) will display phase oscillations due to both the natural qubit energy splitting and the added influence of the dispersive resonator coupling. After extracting the components of the joint Schrödinger equation $i\hbar\partial_t|\Psi\rangle = \hat{H}|\Psi\rangle$ by differentiating Eq. (36), and a bit of algebra, it is straightforward to derive the evolution equation for the coherence directly from differentiating Eq. (38)

$$\partial_t\rho_{01}(t) = i[\omega_q + 2\chi n_w(t)]\rho_{01}(t). \tag{39}$$

Notably, the (complex) weak value of the resonator population naturally appears

$$n_w(t) = \frac{\langle\psi_1(t)|\hat{a}^\dagger\hat{a}|\psi_0(t)\rangle}{\langle\psi_1(t)|\psi_0(t)\rangle} = \bar{n}(t) + i\bar{n}_\gamma(t). \tag{40}$$

This weak value can be understood as a *classical dynamical variable* that completely determines the ensemble-averaged *dynamical* influence of the resonator on the reduced qubit state at each point in time t, even in the absence of any explicit preselection or postselection measurements.

The real part $\bar{n}(t)$ of this weak value is the best estimate of the population of the resonator mode if the qubit transitions between its ground and excited states at time t. According to Eq. (39), this weak value produces a shift $2\chi\bar{n}(t)$ of the natural qubit frequency, commonly known as the *ac Stark shift* [103–105]. Notably, this shift does not involve either of the average mode populations $n_0 = \langle\psi_0|\hat{a}^\dagger\hat{a}|\psi_0\rangle$ or $n_1 = \langle\psi_1|\hat{a}^\dagger\hat{a}|\psi_1\rangle$ that one might naively expect, since the qubit coherence does not pertain to a definite population. Any imaginary part $\bar{n}_\gamma(t)$ of the weak value does not contribute to the ac Stark shift in Eq. (39), but instead produces a decay of the qubit coherence that indicates the resonator coupling is *dephasing* the qubit at a rate $2\chi\bar{n}_\gamma(t)$.

For an explicit example of this reduced evolution, to a good approximation [106–108] a one-sided resonator with energy-decay rate κ (as shown in Fig. 4) that is pumped with a coherent state ε detuned by Δ from the bare resonator frequency will reach a steady state that approximates the pure-state ansatz of Eq. (36), with $|\psi_0\rangle$ and $|\psi_1\rangle$ approximating coherent states with complex classical amplitudes

$$\psi_0 \equiv \langle\psi_0|\hat{a}|\psi_0\rangle = \frac{2\varepsilon}{\kappa}\frac{1}{1 + i2(\Delta - \chi)/\kappa}, \tag{41}$$

$$\psi_1 \equiv \langle\psi_1|\hat{a}|\psi_1\rangle = \frac{2\varepsilon}{\kappa}\frac{1}{1 + i2(\Delta + \chi)/\kappa}. \tag{42}$$

The weak value of the resonator population in Eq. (40) correspondingly has the simple steady-state form (assuming a wide-bandwidth resonator for brevity, with $\chi, \Delta \ll \kappa$)

$$n_w = \psi_1^*\psi_0 \approx \frac{4\varepsilon^2}{\kappa^2}\left[1 + i\frac{4\chi}{\kappa}\right], \tag{43}$$

which produces the ac Stark shift $2\chi\bar{n}$ of the qubit frequency with $\bar{n} = 4\varepsilon^2/\kappa^2$, as well as the dephasing rate $\Gamma = 2\chi\bar{n}_\gamma = 8\chi^2\bar{n}/\kappa$. These expressions for the ac Stark shift and ensemble-average dephasing rate agree with the known results for dispersive qubit measurements in circuit quantum electrodynamics (cQED) [106–109].

We emphasize that the complex weak value of the resonator population in Eq. (40) is *physically* the relevant classical dynamical variable that controls the behavior of the ensemble-averaged reduced qubit evolution in Eq. (39). Indeed, the ac Stark shift of the qubit frequency is the primary method

used in cQED for extracting the average population $\bar{n}$ in the resonator at steady state (e.g., [110]), but we see here that in actuality such a dynamical method probes the *weak value* of that population, and not the population associated with any particular qubit state, or even the total average population $\bar{n}_{\rm tot} = \langle\Psi|\hat{a}^\dagger\hat{a}|\Psi\rangle$. The weak value arises from the interference between the two fields $|\psi_0\rangle$ and $|\psi_1\rangle$ in the resonator that are correlated with the two definite qubit states, on average.

At each time t the real part of n_w indicates the best estimate of the average resonator excitation number seen by a qubit that is not in a definite energy eigenstate, while its imaginary part indicates the best estimate of the backaction on the qubit dynamics caused by the fluctuations of the resonator population around that average value. These latter fluctuations result in ensemble-average dephasing of the reduced qubit state, reinforcing the observation made for weak measurements that a weak value can only describes the average (i.e., classical mean-field) state of affairs for an ensemble of realizations [32]. Indeed, when the time-dependent leakage from the resonator is accounted for, the qubit will not dephase in this manner, but will instead follow a pure *quantum trajectory* [67,107,109,111] that depends on the leakage record. As such, we must necessarily interpret the weak value n_w here as implicitly averaging over many such realizations in practice to produce a classical dynamical variable associated with the classical mean field in the resonator.

IV. DISTURBANCE, QUASIPROBABILITIES, AND HAMILTON-JACOBI

Given the consistent role of the real part of a weak value as a best conditioned observable estimate, we can now observe an intriguing logical tension inherent to the weak value. On one hand, any classical conditioned average must include disturbance to obtain anomalous values [14,15,17,23,24]: the larger the disturbance, the more strange the average can become. On the other hand, the strangeness of the conditioned average in Eq. (33) is greatest when the quantum state is *least* disturbed by an intermediate measurement [93], and even persists when there is no added disturbance to the natural dynamical evolution in the example of Eq. (39).

These two statements imply that any classical (hidden-variable) explanation of a strange weak value as a disturbance effect must satisfy one of two properties: either (a) the quantum state must be a *subjective* (epistemic) quantity that is completely insensitive to whatever physical (ontic) disturbance is occurring, or (b) the relevant disturbance occurs entirely during the *postselection*, and not the intermediate measurement [14].

Classical fields that produce strange weak values in weak measurement experiments satisfy this second property, where the disturbance at the postselection filter causes interference between previously independent field components [27,30–32]. However, quantum systems can display similar interference without wavelike intensities [16,18,95], and permit additional entanglement effects [17] such as the dynamical evolution involving weak values emphasized in the previous section. Especially for such a dynamical role of the weak value, it does not seem possible to ascribe strange weak values to classical disturbance mechanisms without also demanding that the quantum state is itself a fundamentally subjective collection of some more physical microstates; such a demand, while not impossible, must contend with the Pusey-Barrett-Rudolph theorem [112] and the other no-go theorems (reviewed in [113]) for states of such epistemic character.

A. Terletsky-Margenau-Hill and Kirkwood-Dirac

To quantify this logical tension, we can express weak values in a more established and familiar way by rewriting Eq. (1) using the spectral expansion $\hat{A} = \sum_a a|a\rangle\langle a|$ to find $A_w(t) = \sum_a a\,\tilde{p}_{a|i,f}$, where

$$\tilde{p}_{a|i,f} = \frac{\mathrm{Re}\langle f'|a\rangle\langle a|i'\rangle\langle i'|f'\rangle}{|\langle f'|i'\rangle|^2}. \tag{44}$$

This is a conditional *quasiprobability* distribution that weights the eigenvalues of $\hat{A}$ in A_w, and satisfies the normalization $\sum_a \tilde{p}_{a|i,f} = 1$. As a result, if a strange weak value $|A_w| > ||\hat{A}||$ is estimated, then at least one quasiprobability must be negative: $\tilde{p}_{a|i,f} < 0$.

Since the conditioning denominator of Eq. (44) is positive definite, we infer that the joint quasiprobability

$$\tilde{p}_{a,f|i} = \mathrm{Re}\langle f'|a\rangle\langle a|i'\rangle\langle i'|f'\rangle \tag{45}$$

in the numerator of Eq. (44) [also appearing directly in the experimental partial average of Eq. (31)] must be negative. This joint quasiprobability distribution is precisely the *Terletsky-Margenau-Hill distribution* [39–42] that has been used since the late 1930s.

Interestingly, the Terletsky-Margenau-Hill distribution is the real part of the complex quasiprobability distribution introduced even earlier by Kirkwood [29,33–38] as an alternative to the Wigner distribution [43]. This Kirkwood distribution is also known as the *standard-ordering distribution* for quantum phase space [114]. In fact, Dirac later considered this distribution specifically to discuss the classical-to-quantum transition [34], observing that the negativity arises directly from the usual operator noncommutativity of quantum mechanics. Notably, the fully complex weak value that appears in reduced state dynamics is nothing more than a conditioned version of this complex Kirkwood-Dirac quasiprobability distribution.

An important feature of the Kirkwood-Dirac distribution that has recently come to light [29,36] is that any quantum state can be written in an operator basis such that this distribution forms its components. That is, if we define a suitable operator basis

$$\Gamma_{a,f} = \frac{|a\rangle\langle f|}{\langle f|a\rangle}, \tag{46}$$

then we can write any quantum state using the expansion

$$\hat{\rho} = \sum_{a,f}\langle f|a\rangle\langle a|\hat{\rho}|f\rangle\Gamma_{a,f}. \tag{47}$$

As such, the quasiprobabilities $\rho(f,a) = \langle f|a\rangle\langle a|\hat{\rho}|f\rangle$ of the complex Kirkwood-Dirac distribution are a complete quantum state representation for an arbitrary density matrix that is analogous to a complex wave function $\psi(x) = \langle x|\psi\rangle$ for a pure state. Unlike the usual wave function, however, the Kirkwood-Dirac distribution is directly compatible with Bayes' theorem [as used in Eq. (44)], and thus behaves

like a true probability distribution. This notable feature has enabled alternative methods of quantum state tomography by directly measuring complex weak values using von Neumann interactions [29,36,62,115–118], and also permits fully Bayesian quasiprobabilistic reformulations of coherent quantum dynamics [37,38].

B. Wigner distribution and negativity

Importantly, the negativity in such a quasiprobability representation of a quantum state is closely associated to traditional measures of nonclassicality [47,48]. The usual examples of this criterion for nonclassicality are the negativity of the Wigner distribution [43], or the Glauber-Sudarshan P distribution [45,46]; however, the Terletsky-Margenau-Hill distribution in Eq. (45) has the same feature [42].

To emphasize this point, we can in fact relate weak values directly to the Wigner distribution when we are considering infinite-dimensional systems. To see this, consider the Wigner distribution for an initially pure state $|i\rangle$ [43]

$$W_i(x,p) = \int_{-\infty}^{\infty} \langle x - \frac{y}{2}|i\rangle\langle i|x + \frac{y}{2}\rangle e^{ipy/\hbar} \frac{dy}{2\pi\hbar}, \tag{48}$$

where x and p represent the usual classical position and momentum variables. Now suppose we compute the partial average of the momentum p for a fixed x using this joint quasiprobability distribution

$$\begin{aligned}
&\int_{-\infty}^{\infty} p W_i(x,p) dp \\
&= \iint_{-\infty}^{\infty} \langle x - \frac{y}{2}|i\rangle\langle i|x + \frac{y}{2}\rangle (-i\hbar\partial_y e^{ipy/\hbar}) \frac{dy\,dp}{2\pi\hbar}, \\
&= \int_{-\infty}^{\infty} i\hbar\partial_y \left[\langle x - \frac{y}{2}|i\rangle\langle i|x + \frac{y}{2}\rangle\right] \int_{-\infty}^{\infty} e^{ipy/\hbar} \frac{dp}{2\pi\hbar} dy, \\
&= \int_{-\infty}^{\infty} \frac{(-i\hbar\partial_x \langle x - \frac{y}{2}|i\rangle)\langle i|x + \frac{y}{2}\rangle + \text{c.c.}}{2} \delta(y) dy, \\
&= \text{Re}(-i\hbar\partial_x \langle x|i\rangle)\langle i|x\rangle, \\
&= \text{Re}\langle x|\hat{p}|i\rangle\langle i|x\rangle, \\
&= \int_{-\infty}^{\infty} p\, \text{Re}\langle x|p\rangle\langle p|i\rangle\langle i|x\rangle dp.
\end{aligned} \tag{49}$$

In other words, after performing integration by parts the partial average of the Wigner function yields precisely the same result as averaging p over the Terletsky-Margenau-Hill distribution of Eq. (45). It follows that conditioning this partial average directly produces a momentum weak value as the proper local average of momentum at the position x

$$\frac{\int_{-\infty}^{\infty} p W_i(x,p) dp}{\int_{-\infty}^{\infty} W_i(x,p) dp} = \text{Re}\frac{\langle x|\hat{p}|i\rangle}{\langle x|i\rangle}. \tag{50}$$

As with Eq. (44), the denominator is a positive-definite probability, so only the partial average of Eq. (49) in the numerator of Eq. (50) can produce nonclassical behavior.

The inescapable conclusion is that weak values are intimately related to quasiprobability distributions for the quantum formalism in a fundamental and unavoidable way [119]. Moreover, these conditioned averages consistently confirm the best estimate that we expected from the derivation of Eq. (1). Similar relationships can be established for the other quantum phase-space Moyal distributions [44], of which the Wigner function and the Glauber-Sudarshan P distribution [45,46] are special cases. (See [120] for an example that emphasizes this point using nonsymmetrized correlation functions.)

We therefore have the following observation: if strange conditioned averages approximate the functional dependence of the weak value in Eq. (1), then classical hidden variable models will be unable to satisfactorily explain that dependence. If one could, then it would also be able to reproduce other nonclassical statistical features of the quantum theory that arise from the negativity of these quasiprobability distributions (see also [23,51]). Such negativity, in turn, arises from intrinsic quantum interference that is not present in classical systems of particles.

C. Hamilton-Jacobi formalism

As a last, albeit poignant, illustration of the pervasive theoretical role of weak values in the quantum mechanical formalism, let us reexamine the Schrödinger equation for a nonrelativistic particle

$$i\hbar\partial_t|\psi\rangle = \left[\frac{\hat{p}^2}{2m} + V(\hat{x})\right]|\psi\rangle. \tag{51}$$

We can expand this equation into a wave-function form that is local in the coordinate x by contracting it with $\langle x|$ and rearranging

$$\partial_t(i\hbar\,\ln\langle x|\psi\rangle) = \frac{1}{2m}\frac{\langle x|\hat{p}^2|\psi\rangle}{\langle x|\psi\rangle} + V(x). \tag{52}$$

This local form automatically involves the weak values of the kinetic and potential energies.

Noting the familiar structure of this equation, we can then define a complex version of *Hamilton's principle function* (i.e., the classical action) as

$$S(x,t) \equiv -i\hbar\,\ln\langle x|\psi\rangle, \tag{53}$$

with real and imaginary parts

$$\text{Re}S(x,t) = \hbar\Phi(x,t), \tag{54}$$

$$\text{Im}S(x,t) = -\frac{\hbar}{2}\ln\rho(x,t), \tag{55}$$

that naturally isolate the phase (i.e., *eikonal*) Φ and probability density ρ components of the wave function $\langle x|\psi\rangle = \sqrt{\rho(x,t)}\,\exp[i\Phi(x,t)]$.

From this complex action, we can define the classically *conjugate momentum field* $p(x,t)$ that corresponds to the local coordinate x in the usual way

$$p(x,t) = \partial_x S(x,t) = -i\hbar\partial_x\,\ln\langle x|\psi\rangle = \frac{\langle x|\hat{p}|\psi\rangle}{\langle x|\psi\rangle}, \tag{56}$$

which naturally produces a complex momentum weak value as the appropriate classical dynamical variable for the local momentum field. The real part of this weak value is the *Bohmian momentum* [28,121–123]

$$\text{Re}p(x,t) = \hbar\partial_x\Phi(x,t), \tag{57}$$

while the imaginary part is the *osmotic momentum* [124,125]

$$\mathrm{Im}p(x,t) = -\frac{\hbar}{2}\partial_x \ln \rho(x,t), \tag{58}$$

that indicates the logarithmic change of the probability distribution $\rho(x,t) = |\langle x|\psi\rangle|^2$ as x is varied [61]. These local momentum fields jointly act as the average fluidlike momentum that corresponds to a fluctuating classical mean field for the quantum particle [126,127].

Using this complex action, we can thus identically rewrite Schrödinger's Eq. (51) as a straightforward *Hamilton-Jacobi* equation for the classical mean-field x and $p(x,t)$ dynamical variables [128]

$$0 = \partial_t S(x,t) + H[x, p(x,t), t], \tag{59}$$

where the classical Hamiltonian function has the form

$$H[x, p(x,t), t] = \frac{1}{2m}\frac{\langle x|\hat{p}^2|\psi\rangle}{\langle x|\psi\rangle} + V(x), \tag{60}$$

involving local position weak values of the kinetic and potential energies.

Splitting the complex Hamilton-Jacobi equation of Eq. (59) into its independent real and imaginary parts produces

$$0 = \hbar\partial_t \Phi(x,t) + \frac{[\mathrm{Re}p(x,t)]^2}{2m} + V(x) + Q(x), \tag{61}$$

$$0 = \partial_t \rho(x,t) + \partial_x\left(\rho(x,t)\frac{p(x,t)}{m}\right). \tag{62}$$

Note that Eq. (62) is nothing more than the usual continuity equation for the probability density $\rho(x,t)$ in terms of the local Bohmian velocity field $v(x,t) = p(x,t)/m$, while Eq. (61) is the Bohmian Hamilton-Jacobi equation that contains the local kinetic energy, potential energy, and what is usually known as the *quantum potential energy*

$$Q(x,t) \equiv \frac{\mathrm{Var}[\mathrm{Re}p(x,t)]}{2m} = -\frac{\hbar^2}{2m}\frac{\partial_x^2\sqrt{\rho(x,t)}}{\sqrt{\rho(x,t)}}, \tag{63}$$

which is proportional to the *weak momentum variance*

$$\begin{aligned}\mathrm{Var}[\mathrm{Re}p(x,t)] &= \mathrm{Re}\frac{\langle x|[\hat{p} - \mathrm{Re}p(x,t)]^2|\psi\rangle}{\langle x|\psi\rangle}, \\ &= \mathrm{Re}\frac{\langle x|\hat{p}^2|\psi\rangle}{\langle x|\psi\rangle} - \left[\mathrm{Re}\frac{\langle x|\hat{p}|\psi\rangle}{\langle x|\psi\rangle}\right]^2.\end{aligned} \tag{64}$$

The classical particle trajectory limit corresponds to the *eikonal approximation*, or *ray approximation* (familiar from geometrical optics), where we can neglect rapid spatial changes in the probability density $\rho(x,t)$ that arise from the wavelike quantum interference. This approximation allows us to neglect the spatial derivatives of $\rho(x,t)$ that appear in the weak momentum variance of Eq. (64) [or quantum potential of Eq. (63)], as well as in the osmotic momentum of Eq. (58), implying that the local Bohmian momentum field in Eq. (57) has precise and nonfluctuating values at each definite position. In this limit, Eq. (59) becomes purely real and identified with the usual classical Hamilton-Jacobi equation, while the momentum weak value of Eq. (56) reduces to the purely real Bohmian momentum and becomes identified with the usual classical momentum. Similarly, Eq. (62) becomes trivially identified with the usual continuity equation for families of classical trajectories in phase space.

Evidently, at every stage of the usual quantum-classical correspondence through the Hamilton-Jacobi formalism, weak values directly produce the proper classical mean-field dynamical variables, so are the appropriate conditional estimates akin to the expectation values found in the Ehrenfest theorem. This fundamental role of the weak value corroborates the various interpretations we have encountered in this paper, and makes it clear that weak values are an inextricable feature of the quantum formalism that directly pertains to the classical mean-field limit [32].

V. CONCLUSION

Quantum weak values have endured a controversial history, despite the fact that they are essential to the quantum formalism, with their seeming strangeness being a direct consequence of quantum interference. As estimates of observable averages within a bracketed time window, or as classical dynamical variables for the mean-field evolution, their potential for having anomalous values reflects the nonclassicality of the quantum probabilistic structure, and is equivalent to the need for negative quasiprobabilities. This negativity fundamentally arises as an interference effect, both for the classical mean fields and for manifestly quantum systems that permit discrete detection events and entanglement. No classical model can faithfully reproduce the functional dependence of weak value anomalies without also simulating the corresponding features of quantum mechanics.

ACKNOWLEDGMENTS

The author is grateful for discussions with A. N. Jordan, P. B. Dixon, F. Nori, and E. A. Sete. The research was funded by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Office (ARO) Grant No. W911NF-10-1-0334, and also supported from the ARO MURI Grant No. W911NF-11-1-0268.

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100, Phys. Rev. Lett. **60**, 1351 (1988).

[2] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Colloquium: Understanding quantum weak values: Basics and applications, Rev. Mod. Phys. **86**, 307 (2014).

[3] A. G. Kofman, S. Ashhab, and F. Nori, Nonperturbative theory of weak pre- and post-selected measurements, Phys. Rep. **520**, 43 (2012).

[4] Y. Aharonov and L. Vaidman, The two-state vector formalism: an updated review, Lect. Notes Phys. **734**, 399 (2008).

[5] Y. Aharonov, S. Popescu, and J. Tollaksen, A time-symmetric formulation of quantum mechanics, Phys. Today **63**, 27 (2010).

[6] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, The sense in which a "weak measurement" of a spin-1/2 particle's spin component yields a value 100, Phys. Rev. D **40**, 2112 (1989).

[7] C. Ferrie and J. Combes, How the result of a single coin toss can turn out to be 100 heads, Phys. Rev. Lett. **113**, 120404 (2014).

[8] L. Vaidman, Comment on "How the result of a single coin toss can turn out to be 100 heads", arXiv:1409.5386.

[9] E. Cohen, A comment on "How the result of a single coin toss can turn out to be 100 heads", arXiv:1409.8555.

[10] Y. Aharonov and D. Rohrlich, Comment on "How the result of a single coin toss can turn out to be 100 heads", arXiv:1410.0381.

[11] D. Sokolovski, A comment on the paper "How can a result of a single coin toss turn out to be 100 heads" by C. Ferrie and J. Combes, arXiv:1410.0570.

[12] A. Brodutch, Comment on: How the result of a single coin toss can turn out to be 100 heads (arXiv:1403.2362); arXiv:1410.8510.

[13] J. Dressel, S. Agarwal, and A. N. Jordan, Contextual values of observables in quantum measurements, Phys. Rev. Lett. **104**, 240401 (2010).

[14] J. Dressel and A. N. Jordan, Contextual-value approach to the generalized measurement of observables, Phys. Rev. A **85**, 022123 (2012).

[15] N. S. Williams and A. N. Jordan, Weak values and the Leggett-Garg inequality in solid-state qubits, Phys. Rev. Lett. **100**, 026804 (2008).

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

[17] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan, Experimental violation of two-party leggett-garg inequalities with semiweak measurements, Phys. Rev. Lett. **106**, 040402 (2011).

[18] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, Partial-measurement backaction and nonclassical weak values in a superconducting circuit, Phys. Rev. Lett. **111**, 090506 (2013).

[19] J. Dressel and A. N. Korotkov, Avoiding loopholes with hybrid Bell-Leggett-Garg inequalities, Phys. Rev. A **89**, 012125 (2014).

[20] A. J. Leggett and A. Garg, Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? Phys. Rev. Lett. **54**, 857 (1985).

[21] A. J. Leggett, Testing the limits of quantum mechanics: Motivation, state of play, prospects, J. Phys.: Condens. Matter **14**, R415 (2002).

[22] C. Emary, N. Lambert, and F. Nori, Leggett-Garg inequalities, Rep. Prog. Phys. **77**, 016001 (2014).

[23] J. Tollaksen, Pre- and post-selection, weak values and contextuality, J. Phys. A: Math. Theor. **40**, 9033 (2007).

[24] A. C. Ipsen, Disturbance in weak measurements and the difference between quantum and classical weak values, arXiv:1409.3538.

[25] Y. Aharonov, F. Columbo, I. Sabadini, D. C. Struppa, and J. Tollaksen, Some mathematical properties of superoscillations, J. Phys. A: Math. Theor. **44**, 365304 (2011).

[26] M. V. Berry and P. Shukla, Pointer supershifts and superoscillations in weak measurements, J. Phys. A: Math. Theor. **45**, 015301 (2012).

[27] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Realization of a measurement of a "Weak Value", Phys. Rev. Lett. **66**, 1107 (1991).

[28] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg, Observing the average trajectories of single photons in a two-slit interferometer, Science **332**, 1170 (2011).

[29] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, Direct measurement of the quantum wavefunction, Nature (London) **474**, 188 (2011).

[30] K. Y. Bliokh, A. Y. Bekshaev, A. G. Kofman, and F. Nori, Photon trajectories, anomalous velocities, and weak measurements: A classical interpretation, New J. Phys. **15**, 073022 (2013).

[31] K. Y. Bliokh, A. Y. Bekshaev, and F. Nori, Extraordinary momentum and spin in evanescent waves, Nat. Commun. **5**, 3300 (2014).

[32] J. Dressel, K. Y. Bliokh, and F. Nori, Classical field approach to quantum weak measurements, Phys. Rev. Lett. **112**, 110407 (2014).

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

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

[35] S. Chaturvedi, E. Ercolessi, G. Marmo, G. Morandi, N. Mukunda, and R. Simon, Wigner-Weyl correspondence in quantum mechanics for continuous and discrete systems—a Dirac-inspired view, J. Phys. A: Math. Gen. **39**, 1405 (2006).

[36] J. S. Lundeen and C. Bamber, Procedure for direct measurement of general quantum states using weak measurement, Phys. Rev. Lett. **108**, 070402 (2012).

[37] H. F. Hofmann, Complex joint probabilities as expressions of reversible transformations in quantum mechanics, New J. Phys. **14**, 043031 (2012).

[38] C. Bamber and J. S. Lundeen, Observing dirac's classical phase space analog to the quantum state, Phys. Rev. Lett. **112**, 070405 (2014).

[39] Y. P. Terletsky, The limiting transition from quantum to classical mechanics, JETP **7**, 1290 (1937).

[40] H. Margenau and R. N. Hill, Correlation between measurements in quantum theory, Prog. Theor. Phys. **26**, 722 (1961).

[41] L. M. Johansen, Nonclassical properties of coherent states, Phys. Lett. A **329**, 184 (2004).

[42] L. M. Johansen and A. Luis, Nonclassicality in weak measurements, Phys. Rev. A **70**, 052115 (2004).

[43] E. P. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. **40**, 749 (1932).

[44] J. E. Moyal, Quantum mechanics as a statistical theory, Math. Proc. Cambr. Philos. Soc. **45**, 99 (1949).

[45] R. J. Glauber, Photon Correlations, Phys. Rev. Lett. **10**, 84 (1963).

[46] E. C. G. Sudarshan, Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams, Phys. Rev. Lett. **10**, 277 (1963).

[47] R. W. Spekkens, Negativity and contextuality are equivalent notions of nonclassicality, Phys. Rev. Lett. **101**, 020401 (2008).

[48] C. Ferrie, Quasi-probability representations of quantum theory with applications to quantum information science, Rep. Prog. Phys. **74**, 116001 (2011).

[49] S. Kochen and E. P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. **17**, 59 (1967).

[50] N. D. Mermin, Hidden variables and the two theorems of John Bell, Rev. Mod. Phys. **65**, 803 (1993).

[51] M. F. Pusey, Anomalous weak values are proofs of contextuality, Phys. Rev. Lett. **113**, 200401 (2014).

[52] S. Watanabe, Symmetry and physical laws. Part III. Prediction and retrodiction, Rev. Mod. Phys. **27**, 179 (1955).

[53] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Time symmetry in the quantum process of measurement, Phys. Rev **134**, B1410 (1964).

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

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

[56] Y. Aharonov and A. Botero, Quantum averages of weak values, Phys. Rev. A **72**, 052111 (2005).

[57] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, Multiple-time states and multiple-time measurements in quantum mechanics, Phys. Rev. A **79**, 052110 (2009).

[58] M. J. W. Hall, Exact uncertainty relations, Phys. Rev. A **64**, 052103 (2001).

[59] L. M. Johansen, What is the value of an observable between pre- and postselection? Phys. Lett. A **322**, 298 (2004).

[60] M. J. W. Hall, Prior information: How to circumvent the standard joint-measurement uncertainty relation, Phys. Rev. A **69**, 052113 (2004).

[61] J. Dressel and A. N. Jordan, Significance of the imaginary part of the weak value, Phys. Rev. A **85**, 012107 (2012).

[62] H. F. Hofmann, Sequential measurements of non-commuting observables with quantum controlled interactions, New J. Phys. **16**, 063056 (2014).

[63] P. Busch, P. Lahti, and R. F. Werner, Proof of heisenbergs error-disturbance relation, Phys. Rev. Lett. **111**, 160405 (2013).

[64] J. Dressel and F. Nori, Certainty in Heisenberg's uncertainty principle: Revisiting definitions for estimation errors and disturbance, Phys. Rev. A **89**, 022106 (2014).

[65] P. Busch, P. Lahti, and R. F. Werner, Colloquium: Quantum root-mean-square error and measurement uncertainty relations, Rev. Mod. Phys. **86**, 1261 (2014).

[66] P. Campagne-Ibarcq, L. Bretheau, E. Flurin, A. Auffèves, F. Mallet, and B. Huard, Observing interferences between past and future quantum states in resonance fluorescence, Phys. Rev. Lett. **112**, 180402 (2014).

[67] S. J. Weber, A. Chantasri, J. Dressel, A. N. Jordan, K. W. Murch, and I. Siddiqi, Mapping the optimal route between two quantum states, Nature (London) **511**, 570 (2014).

[68] D. Tan, S. J. Weber, I. Siddiqi, K. Mølmer, and K. W. Murch, Prediction and retrodiction for a continuously monitored superconducting qubit, Phys. Rev. Lett. **114**, 090403 (2015).

[69] T. Rybarczyk, S. Gerlich, B. Peaudecerf, M. Penasa, B. Julsgaard, K. Mølmer, S. Gleyzes, M. Brune, J.-M. Raimond, S. Haroche, and I. Dotsenko, Past quantum state analysis of the photon number evolution in a cavity, arXiv:1409.0958.

[70] M. Tsang, Time-symmetric quantum theory of smoothing, Phys. Rev. Lett. **102**, 250403 (2009).

[71] M. Tsang, Continuous quantum hypothesis testing, Phys. Rev. Lett. **108**, 170502 (2012).

[72] S. Gammelmark, B. Julsgaard, and K. Mølmer, Past quantum states of a monitored system, Phys. Rev. Lett. **111**, 160401 (2013).

[73] S. Gammelmark, K. Mølmer, W. Alt, T. Kampschulte, and D. Meschede, Hidden Markov model of atomic quantum jump dynamics in an optically probed cavity, Phys. Rev. A **89**, 043839 (2014).

[74] H. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2007).

[75] H. M. Wiseman and G. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, UK, 2009).

[76] J. Dressel and A. N. Jordan, Quantum instruments as a foundation for both states and observables, Phys. Rev. A **88**, 022107 (2013).

[77] A. Chantasri, J. Dressel, and A. N. Jordan, Action principle for continuous quantum measurement, Phys. Rev. A **88**, 042110 (2013).

[78] L. Vaidman, Quantum theory and determinism, Quant. Stud.: Math. Found. **1**, 5 (2014).

[79] Y. Aharonov, E. Cohen, D. Grossman, and A. C. Elitzur, Can a future choice affect a past measurement's outcome?, arXiv:1206.6224.

[80] R. P. Feynman, The theory of positrons, Phys. Rev. **76**, 749 (1949).

[81] P. Ehrenfest, Bemerkung über die angenäherte gültigkeit der klassischen mechanik innerhalb der quantenmechanik, Z. Phys. **45**, 455 (1927).

[82] C. R. Shalizi and J. P. Crutchfield, Computational mechanics: Pattern and prediction, structure and simplicity, J. Stat. Phys. **104**, 816 (2001).

[83] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney, Time's barbed arrow: Irreversibility, crypticity, and stored information, Phys. Rev. Lett. **103**, 094101 (2009).

[84] C. J. Ellison, J. R. Mahoney, and J. P. Crutchfield, Prediction, retrodiction, and the amount of information stored in the present, J. Stat. Phys. **136**, 1005 (2009).

[85] J. R. Mahoney, C. J. Ellison, R. G. James, and J. P. Crutchfield, How hidden are hidden processes? A primer on crypticity and entropy convergence, Chaos **21**, 037112 (2011).

[86] J. S. Simonoff, *Smoothing Methods in Statistics*, 2nd ed. (Springer-Verlag, New York, 1998).

[87] G. A. Einicke, *Smoothing, Filtering and Prediction: Estimating the Past, Present, and Future* (InTech, Rijeka, 2012).

[88] M. S. Leifer and R. W. Spekkens, Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference, Phys. Rev. A **88**, 052130 (2013).

[89] D. Pegg, S. Barnett, and J. Jeffers, Quantum theory of preparation and measurement, J. Mod. Opt. **49**, 913 (2002).

[90] D. T. Pegg, S. M. Barnett, and J. Jeffers, Quantum retrodiction in open systems, Phys. Rev. A **66**, 022106 (2002).

[91] H. M. Wiseman, Weak values, quantum trajectories, and the cavity-QED experiment on wave-particle correlation, Phys. Rev. A **65**, 032111 (2002).

[92] B. Coecke and R. W. Spekkens, Picturing classical and quantum Bayesian inference, Synthese **186**, 651 (2012).

[93] J. Dressel and A. N. Jordan, Weak values are universal in von neumann measurements, Phys. Rev. Lett. **109**, 230402 (2012).

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

[95] G. J. Pryde, J. L. O'Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, Measurement of quantum weak values of photon polarization, Phys. Rev. Lett. **94**, 220405 (2005).

[96] A. Di Lorenzo, Full counting statistics of weak measurement, Phys. Rev. A **85**, 032106 (2012).

[97] A. M. Steinberg, How much time does a tunneling particle spend in the barrier region? Phys. Rev. Lett. **74**, 2405 (1995).

[98] A. M. Steinberg, Conditional probabilities in quantum theory and the tunneling-time controversy, Phys. Rev. A **52**, 32 (1995).

[99] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

[100] P. Jordan, J. von Neumann, and E. Wigner, On an algebraic generalization of the quantum mechanical formalism, Ann. Math. **35**, 29 (1934).

[101] J. Dressel and A. N. Jordan, Sufficient conditions for uniqueness of the weak value, J. Phys. A: Math. Theor. **45**, 015304 (2012).

[102] J. Dressel, Y. Choi, and A. N. Jordan, Measuring which-path information with coupled electronic Mach-Zehnder interferometers, Phys. Rev. B **85**, 045320 (2012).

[103] A. M. Bonch-Bruevich, N. N. Kostin, V. A. Khodovoĭ, and V. V. Khromov, Changes in the atomic absorption spectrum in the field of a light wave I, Sov. Phys. JETP **29**, 82 (1969).

[104] M. Brune, P. Nussenzveig, F. Schmidt-Kaler, F. Bernardot, A. Maali, J. M. Raimond, and S. Haroche, From lamb shift to light shifts: Vacuum and subphoton cavity fields measured by atomic phase sensitive detection, Phys. Rev. Lett. **72**, 3339 (1994).

[105] D. I. Schuster, A. Wallraff, A. Blais, L. Frunzio, R.-S. Huang, J. Majer, S. M. Girvin, and R. J. Schoelkopf, ac Stark shift and dephasing of a superconducting qubit strongly coupled to a cavity field, Phys. Rev. Lett. **94**, 123602 (2005).

[106] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting, Phys. Rev. A **74**, 042318 (2006).

[107] J. Gambetta, A. Blais, M. Boissonneault, A. A. Houck, D. I. Schuster, and S. M. Girvin, Quantum trajectory approach to circuit QED: Quantum jumps and the Zeno effect, Phys. Rev. A **77**, 012112 (2008).

[108] M. Boissonneault, J. M. Gambetta, and A. Blais, Dispersive regime of circuit QED: Photon-dependent qubit dephasing and relaxation rates, Phys. Rev. A **79**, 013819 (2009).

[109] A. N. Korotkov, in Quantum Bayesian Approach to Circuit QED Measurement, edited by M. Devoret, B. Huard, R. Schoelkopf, and L. F. Cugliandolo, Lecture notes of the Les Houches summer school (Session 96, July 2011) (Oxford University Press, New York, 2011), Chap. 17, p. 533, arXiv:1111.4016.

[110] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis, Fast accurate state measurement with superconducting qubits, Phys. Rev. Lett. **112**, 190504 (2014).

[111] K. W. Murch, S. J. Weber, C. Macklin, and I. Siddiqi, Observing single quantum trajectories of a superconducting quantum bit, Nature (London) **511**, 211 (2013).

[112] M. F. Pusey, J. Barrett, and T. Rudolph, On the reality of the quantum state, Nat. Phys. **8**, 475 (2012).

[113] M. S. Leifer, Is the quantum state real? An extended review of Ψ-ontology theorems, Quanta **3**, 67 (2014).

[114] C. Mehta, Phase-space formulation of the dynamics of canonical variables, J. Math. Phys. **5**, 677 (1964).

[115] J. Z. Salvail, M. Agnew, A. S. Johnson, E. Bolduc, J. Leach, and R. W. Boyd, Full characterisation of polarization states of light via direct measurement, Nat. Photon. **7**, 316 (2013).

[116] M. Malik, M. Mirhosseini, M. P. J. Lavery, J. Leach, M. J. Padgett, and R. W. Boyd, Direct measurement of a 27-dimensional orbital-angular-momentum state vector, Nat. Commun. **5**, 3115 (2014).

[117] G. A. Howland, D. J. Lum, and J. C. Howell, Compressive wavefront sensing with weak values, Opt. Express **22**, 18870 (2014).

[118] M. Mirhosseini, O. S. Magaña-Loaiza, S. M. Hashemi Rafsanjani, and R. W. Boyd, Compressive direct measurement of the quantum wave function, Phys. Rev. Lett. **113**, 090402 (2014).

[119] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes: Quantum Theory for the Perplexed* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005).

[120] A. Bednorz, C. Bruder, B. Reulet, and W. Belzig, Nonsymmetrized correlations in quantum noninvasive measurements, Phys. Rev. Lett. **110**, 250404 (2013).

[121] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables. I, Phys. Rev. **85**, 166 (1952).

[122] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables. II, Phys. Rev. **85**, 180 (1952).

[123] H. M. Wiseman, Grounding Bohmian mechanics in weak values and Bayesianism, New J. Phys. **9**, 165 (2007).

[124] E. Nelson, Derivation of Schrödinger's equation from Newtonian mechanics, Phys. Rev. **150**, 1079 (1966).

[125] D. Bohm and B. J. Hiley, Non-locality and locality in the stochastic interpretation of quantum mechanics, Phys. Rep. **172**, 93 (1989).

[126] E. Madelung, Eine anschauliche deutung der gleichung von Schrödinger, Naturwissenshaften **14**, 1004 (1926).

[127] E. Madelung, Quantentheorie in hydrodynamischer form, Z. Phys. **40**, 322 (1927).

[128] B. J. Hiley and R. Callaghan, Clifford algebras and the Dirac-Bohm quantum-Hamilton-Jacobi equation, Found. Phys. **42**, 192 (2012).