P H Y S I C A L   R E V I E W   L E T T E R S

Contextual Values of Observables in Quantum Measurements

J. Dressel, S. Agarwal, and A. N. Jordan

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA
(Received 23 November 2009; published 15 June 2010)

We introduce contextual values as a generalization of the eigenvalues of an observable that takes into account both the system observable and a general measurement procedure. This technique leads to a natural definition of a general conditioned average that converges uniquely to the quantum weak value in the minimal disturbance limit. As such, we address the controversy in the literature regarding the theoretical consistency of the quantum weak value by providing a more general theoretical framework and giving several examples of how that framework relates to existing experimental and theoretical results.

The weak value (WV) of a quantum operator $\hat{A}$ was introduced in 1988 by Aharonov, Albert, and Vaidman (AAV) [1]. They claimed that if a system is preselected in an initial quantum state $|\psi_i\rangle$, and postselected on a final state $|\psi_f\rangle$, then the result of weakly measuring the operator $\hat{A}$ was not its expectation value, but rather its weak value:

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

In the WV literature, it is customary to point out that for a bounded operator the WV can wildly exceed the eigenvalue range, and, unlike the expectation value, can be complex.

The WV has been both a theoretically and experimentally fruitful concept because of these surprising features. Theoretically, WVs have helped to understand counterintuitive results that involve postselection, even resolving difficulties that arise in quantum mechanics such as Hardy's paradox and apparent superluminal travel [2]. They can be further seen as a single-system test of quantum mechanics: whenever the WV exceeds the eigenvalue range, it rules out certain classes of hidden variable theories and is equivalent to violating a generalized Leggett-Garg inequality [3]. On the experimental side, WVs that exceed the eigenvalue range have been observed in optical systems [4,5]. Indeed, WVs have recently found a practical application, because an anomalously large WV may be used to amplify a small shift in a system parameter. This idea has been exploited in polarization [6] and interferometry [7] experiments to achieve subpicometer position resolution.

In spite of this list of increasingly impressive accomplishments and insights, there is a long history of controversy regarding WVs [8]. Our primary concern here is not the ample philosophical discussion, but rather the physics controversy which we now detail. (i) As defined, the WV (1) is generally complex, so how can it be measured? This was artfully put by Landauer when he asked, "Has anyone seen a stopwatch with complex numbers on its dial?" [9]. (Both the real and imaginary parts of the WV expression

are physically relevant, but are measured with different experiments, see, e.g., [10].) (ii) Another objection is that the AAV formula (1) is not unique, and that other results for the WV can arise with a slightly different measurement setup [11]. (iii) The AAV formula (1) only applies to pure states and arbitrarily weak measurements, so is there a generalization to the cases of mixed states and finite strength measurements? (iv) Perhaps most worrisome, if we view the WV as a weighted average of the eigenvalues of the operator, and the WV exceeds the eigenvalue range, then we can immediately conclude that there must be negative probabilities weighting the eigenvalues. This is dramatically illustrated in the three-box problem [8]. Taken together, these objections are a formidable challenge to the theoretical consistency of the concept of the WV.

The purpose of this Letter is to take the thesis of the second paragraph together with the antithesis of the third and to develop a synthesis. This will be accomplished by going beyond thinking about an observable in terms of its eigenvalues and advancing the concept that the measurement results must be interpreted within their own context. This idea will give rise to the notion of the contextual values (CV) of an observable. The synthesis occurs by showing how the WV can be subsumed as a special case in the CV formalism. This formalism contains additional information that is able to clarify the objections listed above in an organic fashion. In addition, we will derive a generalized WV formula that can be practically applied to a much larger class of problems that involve postselected averages. We also provide illustrative examples and demonstrate that we can recover a number of predictions for postselected averages in the literature that had been calculated for specific situations using other, more involved techniques.

Contextual values.—The idea behind CV stems from the observation that the intrinsically measurable quantities in the quantum theory are the outcome probabilities for a particular measurement setup. In order to calculate averages of an observable from the measured probabilities, it is empirically necessary to assign values to each outcome as a

240401-1

separate step. This secondary step becomes implicit in the operator formalism, so is rarely mentioned explicitly.

Under projective measurements such a value assignment seems trivial: the outcomes are perfectly distinguishable and easily assigned meaningful values, which become the eigenvalues for the observable operator. However, when generalizing the measurement to nonprojective outcomes, one must be careful to correctly match the measured probabilities to a consistent value set. We will now show that the observable operator can be expanded in multiple ways corresponding to the possible measurement strategies and value assignments (4).

To illustrate how the observable values change, we consider a general measurement that is fully characterized by a set of N measurement operators, $\mathfrak{M} = \{\hat{M}_j\}$, which represent the operation of an N-outcome measurement apparatus on the system. This measurement context generates a corresponding positive operator-valued measure (POVM), $\mathfrak{E} = \{\hat{E}_j = \hat{M}_j^\dagger \hat{M}_j\}$, which, in conjunction with an arbitrary density operator $\hat{\rho}$, determines a probability measure, $P_j = \mathrm{Tr}[\hat{E}_j \hat{\rho}]$, on the space of outcomes. Empirically, these probabilities can be measured as the relative frequencies of obtaining each outcome.

Suppose that we wish to reconstruct the average value of an observable $\mathcal{A}$ using this setup. In order to obtain the correct average as a weighted sum,

$$\langle \mathcal{A} \rangle = \sum_j \alpha_j P_i = \sum_j \alpha_j \, \mathrm{Tr}[\hat{E}_j \hat{\rho}], \qquad (2)$$

we must explicitly assign a set of values $\{\alpha_j\}$ to the N outcomes of the apparatus.

We require that such an average be independent of $\mathfrak{M}$, and thus a property of the state $\hat{\rho}$ itself, which constrains the possible value assignments. Specifically, (2) should be equivalent to the average under a projective context, $\mathfrak{P} = \{\hat{\Pi}_k\}$,

$$\langle \mathcal{A} \rangle = \sum_k a_k \, \mathrm{Tr}[\hat{\Pi}_k \hat{\rho}] = \mathrm{Tr}[\hat{A} \, \hat{\rho}], \qquad (3)$$

where we have identified the projection-valued measure (PVM) and the corresponding eigenvalues $\{a_k\}$ as the spectral decomposition of a Hermitian operator $\hat{A}$.

We now assert that the values $\{\alpha_j\}$ should be determined solely by $\mathfrak{M}$ so the equality between (2) and (3) should hold for any state $\hat{\rho}$. It follows that the operator $\hat{A}$ can be expanded in multiple ways corresponding to possible contexts,

$$\hat{A} = \sum_j \alpha_j \hat{E}_j = \sum_k a_k \hat{\Pi}_k. \qquad (4)$$

This operator identity defines the contextual values $\{\alpha_j\}$ of an operator $\hat{A}$ under a compatible measurement context $\mathfrak{M}$ as a generalization of its eigenvalues that allows the full empirical reconstruction of the associated physical observable $\mathcal{A}$.

Indeed we can use this operator equality to obtain the nth moment of $\mathcal{A}$ using the same experimental setup,

$$\langle \mathcal{A}^n \rangle = \sum_{j_1, \dots, j_n} \alpha_{j_1} \dots \alpha_{j_n} \, \mathrm{Tr}[\hat{E}_{j_1} \dots \hat{E}_{j_n} \hat{\rho}], \qquad (5)$$

provided that the N measurement operators in $\mathfrak{M}$ and $\hat{A}$ all commute. In that case the correlation probabilities, $P_{j_1, \dots, j_n} = \mathrm{Tr}[\hat{M}_{j_n} \cdots \hat{M}_{j_1} \hat{\rho} \hat{M}_{j_1}^\dagger \cdots \hat{M}_{j_n}^\dagger]$, appearing in (5) can be measured as the relative frequencies of the N^n outcome sequences of n repeated measurements.

In addition to permitting the reconstruction of the moments of the observable, its CV are of great interest in their own right. Note that the moments of the CV themselves, $\sum_j \alpha_j^n P_i$, only coincide with the moments of $\mathcal{A}$ for $n = 1$. Such moments nevertheless contain important physics about both the context $\mathfrak{M}$ and the observable $\mathcal{A}$, as we consider shortly.

The construction (4), in general, will fail if $N < \dim \hat{A}$, will be unique if $N = \dim \hat{A}$, and will be underspecified if $N > \dim \hat{A}$. The latter case results in an infinite number of possible solutions, $\{\alpha_j\}$. As such, we propose that the physically sensible choice of CV is the least redundant set uniquely related to the eigenvalues through the Moore-Penrose pseudoinverse.

To illustrate the construction of the least redundant set of CV, we consider the case when $\{\hat{M}_j\}$ and $\hat{A}$ all commute. It follows that they can all be diagonalized in the same PVM, and (4) is isomorphic to a matrix equation $\vec{a} = F\vec{\alpha} = (\sum_j F_{kj} \alpha_j)$, with the components of F given by $F_{kj} = \mathrm{Tr}[\hat{\Pi}_k \hat{E}_j]$. To find the pseudoinverse of the matrix F, we find its singular value decomposition, $F = U \Sigma V^T$, where the orthogonal matrices, U and V, are composed of the eigenvectors of FF^T and $F^T F$, respectively, and Σ is a diagonal matrix composed of the singular values of F [12]. The pseudoinverse of F can be constructed as $F^+ = V \Sigma^+ U^T$, where Σ^+ is the diagonal matrix constructed from Σ^T by inverting all nonzero singular values. It then follows that a uniquely specified set of CV can be defined as $\vec{\alpha}_0 = F^+ \vec{a}$. Other solutions of (4) can be written $\vec{\alpha} = \vec{\alpha}_0 + \vec{x}$, where $\vec{x}$ is in the null space of F.

Conditioned averages.—Suppose we now wish to generalize the notion of an observable average by conditioning the average on the result of a second measurement. To do so we weight the CV, $\{\alpha_j^{(1)}\}$, already determined by the first measurement context, $\mathfrak{M}^{(1)}$, with a set of conditional probabilities generated from a second measurement.

A second measurement is performed on the system with a context, $\mathfrak{M}^{(2)} = \{\hat{M}_f^{(2)}\}$, yielding the full context for the (ordered) sequential measurements, $\mathfrak{M}^{(1,2)} = \{\hat{M}_f^{(2)} \hat{M}_j^{(1)}\}$. The corresponding POVM, $\mathfrak{E}^{(1,2)} = \{\hat{E}_{jf}^{(1,2)} = \hat{M}_j^{(1)\dagger} \hat{M}_f^{(2)\dagger} \hat{M}_f^{(2)} \hat{M}_j^{(1)}\}$, gives the measurable probabilities, $P_{jf} = \mathrm{Tr}[\hat{E}_{jf}^{(1,2)} \hat{\rho}]$ and $P_f = \sum_j P_{jf}$. From these probabilities we can define a conditional probability in the usual
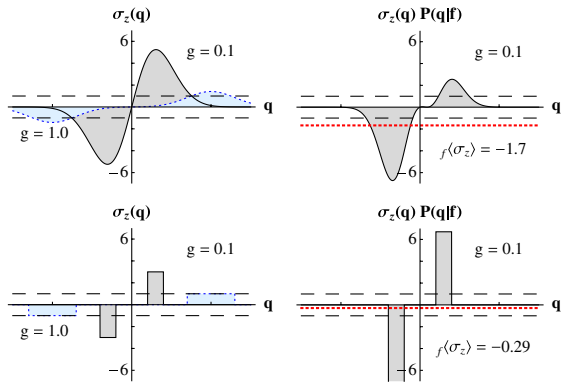
FIG. 1 (color online).   CV (left) and conditioned CV (right) computed using the setup described above Eq. (10) with a coupling strength $g = 0.1$ and a pointer distribution width $\sigma = 0.3$. The preparation state is $\psi(47\pi/32)$, and postselection state is $f(\pi/2)$. Two detector distributions are compared: Gaussian (top) with conditioned average (right, dotted line) violating the eigenvalue range (dashed lines), and box (bottom) showing the effect of the shape on the WV convergence rate. Strong measurement regime CV for each distribution with $g = 1.0$ are also shown for comparison (left, dotted line).

way, $P_{j|f} = P_{jf}/P_f$, and thus define our main result, the conditioned average of an observable,

$$_f\langle \mathcal{A} \rangle = \sum_j \alpha_j^{(1)} P_{j|f} = \frac{\sum_j \alpha_j^{(1)} \mathrm{Tr}[\hat{E}_{jf}^{(1,2)} \hat{\rho}]}{\sum_j \mathrm{Tr}[\hat{E}_{jf}^{(1,2)} \hat{\rho}]}. \quad (6)$$

Unlike $\langle \mathcal{A} \rangle$, $_f\langle \mathcal{A} \rangle$ is explicitly dependent on $\mathfrak{M}^{(1)}$ and the CV. Hence, it must generally be interpreted as encoding information not just about the observable $\mathcal{A}$, but also about the measurement context itself.

Conditioned averages must be bounded by the minimum and maximum CV. However, since the CV range is not the same as the eigenvalue range for the observable, values of $_f\langle \mathcal{A} \rangle$ outside the eigenvalue range may be obtained.

Weak values.—To find the weak limit of (6) we note that any measurement context continuously connected to the identity operation can be decomposed into the form $\mathfrak{M} = \{\hat{U}_j(g)\hat{E}_j^{1/2}(g)\}$, where g is a measurement strength parameter, and $\hat{U}_j(g) = \exp[ig\hat{G}_j]$ are unitary operators generated by Hermitian operators $\hat{G}_j$ by Stone's theorem. Since the POVM elements are bounded and partition unity, they can also be expanded as $\hat{E}_j(g) = p_j\hat{1} + g\hat{E}_j' + \mathcal{O}(g^2)$, where $\sum_j p_j = 1$.

Writing the initial context in (6) to first order in g, we find that as $g \to 0$, the weak limit generally depends explicitly on $\{\hat{G}_j\}$ and $\{\alpha_j\}$, and thus will change depending on how it is measured and how the CV are chosen (see also [11]). However, if $\forall j$, $[\hat{G}_j, \hat{\rho}] = 0$, so the state is minimally disturbed, then the context dependence vanishes and a generalized WV [13] is uniquely defined as the quantity

$$A_w = \mathrm{Tr}[\hat{E}_f^{(2)}\{\hat{A}, \hat{\rho}\}]/2\,\mathrm{Tr}[\hat{E}_f^{(2)}\hat{\rho}], \quad (7)$$

where $\{\cdot, \cdot\}$ denotes the anticommutator. For a pure initial state, $\hat{\rho} = |\psi_i\rangle\langle\psi_i|$, a "pure POVM," $\forall j$, $\hat{U}_j = \hat{1}$, and a strong final measurement, $\hat{E}_f^{(2)} = |\psi_f\rangle\langle\psi_f|$ (as considered by AAV and most subsequent applications), (7) can be written as the real part of (1) [14].

Photon polarization.—As a first example, a photon polarization measurement of tunable strength is detailed in [5] and can be readily interpreted using CV. The Stokes observable $\mathcal{S}_z$ is measured in the horizontal-vertical polarization basis with the context $\hat{M}_+ = \gamma\hat{\Pi}_H + \bar{\gamma}\hat{\Pi}_V$ and $\hat{M}_- = \bar{\gamma}\hat{\Pi}_H + \gamma\hat{\Pi}_V$, leading to the POVM, $\hat{E}_\pm = (1/2)(\hat{1} \pm g\hat{\sigma}_z)$, where $g = \gamma^2 - \bar{\gamma}^2$. The CV for this scenario are uniquely $\pm 1/g$, which exceed the eigenvalue range for $|g| < 1$.

Computing the conditioned average (6) of $\mathcal{S}_z$ with an initial state of $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$ and a final state $|f\rangle = (|H\rangle - |V\rangle)/\sqrt{2}$ yields

$$_f\langle \mathcal{S}_z \rangle = (|\alpha|^2 - |\beta|^2)/(1 - 4\gamma\bar{\gamma}\,\mathrm{Re}[\alpha\beta^*]), \quad (8)$$

which is the result obtained in [5] through direct computation using an ancilla system.

AAV setup.—For the original AAV–von Neumann setup [1], implemented using photons in [4], an interaction Hamiltonian $\hat{H}_{\mathrm{int}} = g\delta(t - t_0)\hat{\sigma}_z \otimes \hat{p}_D$ of a qubit with a detector having momentum operator $\hat{p}_D$ leads to a continuous set of measurement operators on the system $\hat{M}(q) = \langle q - g\hat{\sigma}_z|\Phi_D\rangle$, where $|\Phi_D\rangle$ is the initial detector state, q is the position of the detector pointer, and g is the coupling strength. Hence the POVM is simply the initial detector probability density shifted in position by an operator, $\hat{E}(q) = P_D(q - g\hat{\sigma}_z) = |\langle q - g\hat{\sigma}_z|\Phi_D\rangle|^2$.

Performing the pseudoinversion [12] leads to the least redundant set of CV, $\sigma_z(q)$, for the z component of the pseudospin,

$$\sigma_z(q) = [P_D(q - g) - P_D(q + g)]/(a - b(g)),$$

$$a = \int P_D(q)^2 dq,$$

$$b(g) = \int P_D(q - g)P_D(q + g)dq. \quad (9)$$

From these expressions it is immediately clear that if the shift g is small compared to the width of the initial distribution $P_D(q)$, then the POVM will be nearly proportional to the identity and the measure of overlap, $b(g)$, will be nearly equal to the distribution constant a. Hence, the measurement will be weak and the CV will diverge independently of any specific system states.

For a Gaussian distribution, $P_D(q) = \exp[-q^2/2\sigma^2]/\sigma\sqrt{2\pi}$, $a = 1/2\sigma\sqrt{\pi}$ and $b(g) = a\exp[-(g/\sigma)^2]$. Hence, the CV are $\sigma_z(q) = \sqrt{2}\exp[-q^2/2\sigma^2][\sinh(qg/\sigma^2)/\sinh(g^2/2\sigma^2)]$ [15]. With an initial state $\psi(\alpha) = (\cos(\alpha/2), \sin(\alpha/2))$ in the z basis conditioned on a final state $f(\pi/2) = (2)^{-1/2}(1, 1)$, (6) yields
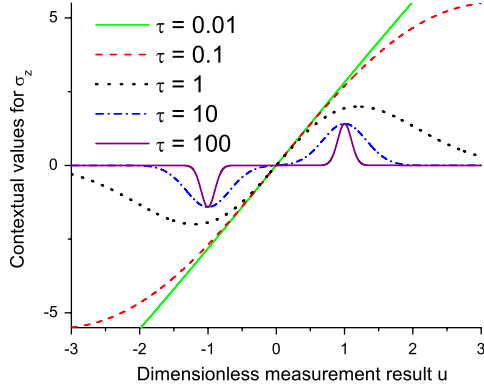
FIG. 2 (color online). QPC CV for a range of measurement times and a fixed coupling constant. For small times $\sigma_z(u) \to 2\sqrt{2}u$, while for long times $\sigma_z(u)$ is only nonzero at ± 1.

$$_f\langle S_z \rangle = \cos\alpha/(1 + \sin\alpha \exp[-g^2/2\sigma^2]). \tag{10}$$

The strong limit as $(g/\sigma) \to \infty$ is $\cos\alpha$, and the weak limit as $(g/\sigma) \to 0$ is the WV $\cot(\alpha/2 + \pi/4)$. The same WV is also obtained for any $P_D(q)$ according to (7), though the speed of convergence depends on the distribution shape. Figure 1 illustrates the effect of the shape by comparing Gaussian and box distributions.

Quantum point contact (QPC) detector.—A generalization of the AAV result to continuous measurements in solid state detectors is given in [3]. The measurement operators are constructed in terms of an average current through a QPC, $\bar{I} = (1/t)\int_0^t I(t')dt'$, and two characteristic currents, I_1 and I_2, that indicate the position of an electron in a nearby double quantum dot. The current detection probabilities are assumed Gaussian with a variance of $\sigma^2 = S_I/2t$, where S_I is the detector shot noise power and t is the averaging time for the current.

The system can be mapped to the AAV case by making the identifications $I_0 = (I_1 + I_2)/2$, $q = \bar{I} - I_0$, $g = (I_1 - I_2)/2$, and $\sigma^2 = S_I/2t$. Since the characteristic currents for the electron position are fixed, the coupling strength g is also fixed. Hence it is useful to consider instead the dimensionless variables, $u = q/g$, and $\tau = (g/\sigma)^2 = t/T_m$, where T_m is the characteristic measurement time. Figure 2 shows a range of CV for varying τ.

The system is started with an initially arbitrary density matrix $\hat{\rho} = [\rho_{ij}]$ written in the z basis. After being measured for the averaging time t, it is then rotated around the $\hat{\sigma}_x$ axis by an angle θ and measured strongly to be in a final state $|f\rangle = |+1\rangle$ along the new z axis. Computing the conditioned average using (6) yields

$$_f\langle S_z \rangle = \frac{\cos^2(\frac{\theta}{2})\rho_{11} - \sin^2(\frac{\theta}{2})\rho_{22}}{\cos^2(\frac{\theta}{2})\rho_{11} + \sin^2(\frac{\theta}{2})\rho_{22} - \sin\theta\,\mathrm{Im}[\rho_{12}]e^{-\tau/2}}, \tag{11}$$

which matches the result in [3] computed using the quantum Bayesian approach [15].

Conclusions.—We have shown how to completely reconstruct an observable using an arbitrary measurement setup by defining contextual values (4), or generalized eigenvalues, that account for the measurement context. The approach allows a conditioned average (6) to be defined that correctly reproduces existing results in the literature and converges uniquely to a generalized weak value (7) in the minimum disturbance limit.

This conditioned average addresses much of the controversy surrounding weak values in the literature, since it (i) is a purely real empirical average, (ii) converges to the real part of the standard weak value (1) uniquely in the minimum disturbance limit with pure initial and final states, (iii) applies to any measurement strength, arbitrary quantum density operators, and POVM postselections, and (iv) is constructed from well-behaved, measurable probabilities.

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988); see also M. Büttiker, Phys. Rev. B **27**, 6178 (1983).

[2] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992); Y. Aharonov *et al.*, Phys. Lett. A **301**, 130 (2002); D. Rohrlich and Y. Aharonov, Phys. Rev. A **66**, 042102 (2002).

[3] A. J. Leggett and A. Garg, Phys. Rev. Lett. **54**, 857 (1985); N. S. Williams and A. N. Jordan, Phys. Rev. Lett. **100**, 026804 (2008).

[4] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, Phys. Rev. Lett. **66**, 1107 (1991); N. Brunner, *et al.*, Phys. Rev. Lett. **93**, 203902 (2004); T. C. Ralph *et al.*, Phys. Rev. A **73**, 012113 (2006).

[5] G. J. Pryde *et al.*, Phys. Rev. Lett. **94**, 220405 (2005).

[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); Phys. Rev. A **80**, 041803(R) (2009).

[8] The WVs controversy is chronicled by Y. Aharonov and L. Vaidman, arXiv:quant-ph/0105101v2.

[9] M. Büttiker (private communication).

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

[11] S. Parrott, arXiv:0909.0295; arXiv:0908.0035

[12] The singular value decomposition can be readily generalized to an operator $\hat{A}$ and/or POVM with infinite dimension as long as the resulting operator $\hat{F}$ is compact.

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

[14] Under strong (PVM) measurements, (6) reduces to $_f\langle A \rangle = (\sum_k |\langle f|k\rangle|^2 a_k |\langle k|\psi\rangle|^2)/(\sum_k |\langle f|k\rangle|^2 |\langle k|\psi\rangle|^2)$, in contrast with the weak limit.

[15] If $P_D(q)$ is symmetric about the mean, $_f\langle S_z \rangle$ becomes independent of the choice of CV. Hence, the generic CV often implied in the literature, $\sigma_z(q) = q/g$, give the same result as the least redundant CV.