Sequential weak measurement

Graeme Mitchison*
Centre for Quantum Computation, DAMTP, University of Cambridge, Cambridge CB3 0WA, United Kingdom

Richard Jozsa†
Department of Computer Science, University of Bristol, Bristol BS8 1UB, United Kingdom

Sandu Popescu‡
H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom
and Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, United Kingdom
(Received 3 August 2007; published 10 December 2007)

The notion of weak measurement provides a formalism for extracting information from a quantum system in the limit of vanishing disturbance to its state. Here we extend this formalism to the measurement of sequences of observables. When these observables do not commute, we may obtain information about joint properties of a quantum system that would be forbidden in the usual strong measurement scenario. As an application, we provide a physically compelling characterization of the notion of counterfactual quantum computation.

DOI: 10.1103/PhysRevA.76.062105 PACS number(s): 03.65.Ta, 03.67.−a, 02.20.Qs

I. INTRODUCTION

Quantum mechanics is still capable of giving us surprises. A good example is the concept of weak measurement discovered by Aharonov and his group [1,2], which challenges one of the canonical dicta of quantum mechanics: that noncommuting observables cannot be simultaneously measured.

Standard measurements yield the eigenvalues of the measured observables, but at the same time they significantly disturb the measured system. In an ideal von Neumann measurement the state of the system after the measurement becomes an eigenstate of the measured observable, no matter what the original state of the system was. On the other hand, by coupling a measuring device to a system weakly it is possible to read out certain information while limiting the disturbance to the system. The situation becomes particularly interesting when one postselects on a particular outcome of the experiment. In this case the eigenvalues of the measured observable are no longer the relevant quantities; rather the measuring device consistently indicates the *weak value* given by the Aharonov-Albert-Vaidman (AAV) formula [2,3]:

$$A_w = \frac{\langle \psi_f | A | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle}, \tag{1}$$

where A is the operator whose value is being ascertained, $|\psi_i\rangle$ is the initial state of the system, and $|\psi_f\rangle$ is the state that is postselected (e.g., by performing a measurement). The significance of this formula is that, if we couple a measuring device whose pointer has position coordinate q to the system $\mathcal{S}$, and subsequently measure q, then the mean value $\langle q \rangle$ of the pointer position is given by

$$\langle q \rangle = g \, \mathrm{Re}[A_w], \tag{2}$$

where Re denotes the real part. This formula requires the initial pointer wave function to be real and of zero mean, but these assumptions will be relaxed later. The coupling interaction is also taken to be the standard von Neumann measurement interaction $H = gAp$. The coupling constant g is assumed to be small, but we can determine A_w to any desired accuracy if enough repeats of the experiment are carried out.

The formula (1) implies that, if the initial state $|\psi_i\rangle$ is an eigenstate of a measurement operator A, then the weak value postconditioned on that eigenstate is the same as the classical (strong) measurement result. When there is a definite outcome, therefore, strong and weak measurements agree. However, weak measurement can yield values outside the normal range of measurement results, e.g., spins of 100 [2]. It can also give complex values, whose imaginary part correspond to the pointer momentum. In fact, the mean of the pointer momentum is given by

$$\langle p \rangle = 2gv \, \mathrm{Im}[A_w], \tag{3}$$

where Im denotes the imaginary part and v is the variance in the initial pointer momentum.

The fact that one hardly disturbs the system in making weak measurements means that one can in principle measure different variables in succession. We follow this idea up in this paper.

II. A PARADOX

Weak measurement has proved to be a valuable tool in analyzing paradoxical quantum situations such as Hardy's paradox [1,4]. To illustrate the idea of sequential weak measurement and its potential applications we first construct a quantum paradox. Consider the *double interferometer*, the optical circuit shown in Fig. 1, where a photon passes through two successive interferometers. This configuration has been considered previously by Bläsi and Hardy [5] in

*g.j.mitchison@damtp.cam.ac.uk
†r.jozsa@bristol.ac.uk
‡s.popescu@bristol.ac.uk

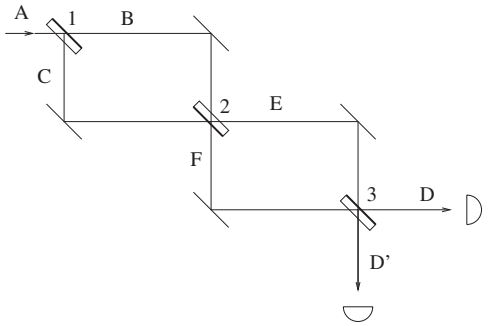

FIG. 1. The double interferometer: an optical circuit in which a photon, injected along path A, passes through two interferometers, represented by paths B and C and paths E and F. Finally, the photon is postselected at the detector D. The beamsplitters are shown with their reflecting surface marked in black.

another context. Using the labels of the paths shown in the figure, and denoting the action of the ith beam splitter by U_i, the system evolves as follows:

$$U_1|A\rangle = (|B\rangle + |C\rangle)/\sqrt{2}, \tag{4}$$

$$U_2|B\rangle = (|E\rangle + |F\rangle)/\sqrt{2}, \quad U_2|C\rangle = (|E\rangle - |F\rangle)/\sqrt{2}, \tag{5}$$

$$U_3|E\rangle = (-|D\rangle + |D'\rangle)/\sqrt{2}, \quad U_3|F\rangle = (|D\rangle + |D'\rangle)/\sqrt{2}. \tag{6}$$

(The signs here are determined by the fact that reflection on the silvered outer surface of a beam splitter gives a phase of π whereas transmission or reflection by the inner surface gives zero phase.)

Suppose now that we select a large number N of successful runs of our experiment, i.e., those runs where the photon is detected by the detector D.

We can now make the following statements about this situation:

(1) *All photons go through path E.* Equations (4) and (5) tell us that if a photon is injected along path A, it must exit the first interferometer along path E. Consequently, if we measure the observable P_E, the projector for path E, we find the total number of photons detected is $N_E=N$ with certainty.

(2) *All photons go through path C.* The second interferometer is arranged in such a way that any photon entering along path B will end up at D'. Hence, a very simple calculation shows that if, instead of measuring N_E, we measure N_C, the number of photons going along path C in all N runs of the experiment, we will obtain with certainty $N_C=N$.

(3) *When photons go through path C, a subsequent measurement reveals that half of them must go through path E and half through path F.* If we measure the position of the photons in the first interferometer and find that all go via C, then a subsequent measurement of N_E and N_F must yield $N/2$ in each case, up to statistical fluctuations. (In fact this is true regardless of whether or not all photons end up eventually at D.)

(4) *When photons go through path E, a subsequent measurement reveals that half of them must have come via path B and half via path C.* This last statement is similar to point (3) above.

The above four statements seem to imply a paradoxical situation. On the one hand, statement (2) tells us, when we pool all the results, that all N photons go via path C; together with statement (3) this implies that the number of photons that go along path E must be $N/2$. On the other hand, statement (1) tells us that all N photons actually go along path E. A similar contradiction arises in connection with the number of photons going along path C. On the one hand, statement (1) tells us that all photons go via E; together with statement (4) this implies that the number of photons that go along path C must be only $N/2$. On the other hand, statement (2) tells us that all N photons actually go along path C.

The usual way of resolving this paradox is to say that the above statements refer to measurements that cannot all be made simultaneously. Indeed, it is true that if we measure P_E we find it is 1 with certainty, but *only* if we do not also measure P_C. If we also measure P_C in the same experiment, then it is no longer the case that $P_E=1$. Similarly, it is true that $P_C=1$ with certainty, but *only* if we do not also measure N_E. If we also measure P_E in the same experiment, then it is no longer the case that $P_E=1$. So, we are told, the statements (1)–(4) above have no simultaneous meaning, for they do not refer to the same experiment. Hence there is no paradox: In formulating the paradox presented above we made use of facts that are not all simultaneously true.

On the other hand, as is emphasized in [3], one should not dismiss such paradoxes too lightly. Indeed it is possible to make a tradeoff: By accepting some imprecision in measuring P_E, P_C, etc., we can limit the disturbance these measurements produce. The way to do this is to weaken the coupling of the measuring devices to the photons.

Since the disturbance is now small, we can make all the measurements in the same experiment, and we expect all the statements (1)–(4) to be true. Hence we expect $N_E=N$, $N_C=N$ and obviously $N_F=0$ and $N_B=0$. On the other hand, we also expect that N_{CE} and N_{CF}, the total numbers of photons that went along C and subsequently along E or F, respectively, should both be equal to $N/2$; this is because all the N photons go via C and half of them should continue along E and half along F. Also we expect N_{CF}, the number of photons that went along C and subsequently along E, to be $N_{CE}=N/2$. Similarly we expect that $N_{CE}$and N_{BE} should both be $N/2$, since all N photons go along E and half of them must come via B and half via C.

While all the above predictions seem reasonable, here is the surprise: Overall we have only N photons. They could have moved along four possible trajectories: BE, BF, CE or CF (see Fig. 2). Since $N_{BE}+N_{BF}+N_{CE}+N_{CF}=N$ and since $N_{BE}=N_{CE}=N_{CF}=N/2$ it must be the case that $N_{BF}=-N/2$. Furthermore, our prediction has a remarkable internal consistency. We know that the total number of photons that go along F must be zero. They can arrive at F in two ways, either by BF or CF. Thus $N_F=N_{BF}+N_{CF}$. As noted above, $N_{CF}=N/2$, but no photons are supposed to go through F. This is due to the fact that N_{BF} is negative, i.e., $N_{BF}=-N/2$.

The above predictions seem totally puzzling, no less puzzling than the original paradox. However, what we have now is not a mere interpretation that can simply be dismissed.

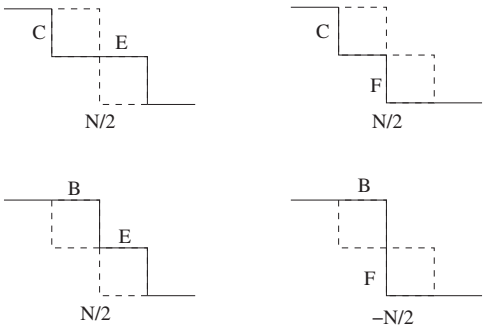

FIG. 2. Paths through the double interferometer, and the number of photons that follow the indicated path. Thus for instance $N_{BE}=N/2$. Note however the curious prediction $N_{BF}=-N/2$.

These are now predictions about the results of real measurements—in particular the weak measurement of the number of photons that passes along path B and then along path F. This is a *two-time* measurement.

In general, by ensuring that the measurement interaction is weak, we can consider *sequences* of measurements. Describing such measurements is the main subject of our paper. In the process, we will formally derive the strange predictions made above for the double interferometer, and will discuss the interpretation of weak measurements. Finally, we apply these ideas to counterfactual computation, which is a catch-all for numerous counterfactual phenomena including, for example, interaction-free measurement [6].

III. SEQUENTIAL WEAK MEASUREMENTS

The situation we shall consider is where a system $\mathcal{S}$ evolves unitarily from an initial state $|\psi_i\rangle$ to a final postselected measurement outcome $\langle\psi_f|$. At various points, observables may be measured weakly. Here we consider the scenario where there is a single copy of the system, with the measuring device weakly coupled to it. Generally, reliable information will only be obtained after many repeats of the given experiment.

In the simplest case where there is just one observable, A, say, we assume the evolution from $|\psi_i\rangle$ to the point where A is measured is given by U, and from this point to the postselection the evolution is given by V. Then we can rewrite Eq. (1) as

$$A_w = \frac{\langle\psi_f|VAU|\psi_i\rangle}{\langle\psi_f|VU|\psi_i\rangle}, \tag{7}$$

and the mean of the pointer is given by Eq. (2) as before.

Consider next the case of two observables, A_1 and A_2, measured at different times on a system $\mathcal{S}$. We assume the system evolves under U from $|\psi_i\rangle$ to the point where A_1 is measured, then under V to the point where A_2 is measured, and finally under W to $|\psi_f\rangle$. Our strategy is to use two measuring devices for measuring A_1 and A_2. Let the positions of their pointers be denoted by q_1 and q_2, respectively. We couple them to the system at successive times, measure q_1 and q_2, and then take the product q_1q_2.

We begin, therefore, with the weak coupling of system and pointers, with the usual von Neumann-type Hamiltonians for measuring A_1 and A_2. The state of system and pointers after this coupling is

$$\Psi_{\mathcal{S}\mathcal{M}_1\mathcal{M}_2} = e^{-igp_2A_2}Ve^{-igp_1A_1}U|\psi_i\rangle_{\mathcal{S}}\phi(q_1)\phi(q_2), \tag{8}$$

where p_1 and p_2 are the two pointer momenta (the label $\mathcal{S}$ refers to the system and $\mathcal{M}_1$, $\mathcal{M}_2$ to the pointers). Here $\phi(q)$ is the initial pointer distribution, and we have assumed, for simplicity, that the two pointers have identical initial distributions and equal coupling constants g. Post-selecting on $\langle\psi_f|$ gives the state of the pointers as

$$\Psi_{\mathcal{M}_1\mathcal{M}_2} = \langle\psi_f|We^{-igp_2A_2}Ve^{-igp_1A_1}U|\psi_i\rangle\phi(q_1)\phi(q_2). \tag{9}$$

As g is small, we can approximate the state as

$$\begin{aligned}\Psi_{\mathcal{M}_1\mathcal{M}_2} = \langle\psi_f|\Bigg[W\left(1 - igp_2A_2 - \frac{g^2}{2}p_2^2A_2^2 + \cdots\right) \\ \times V\left(1 - igp_1A_1 - \frac{g^2}{2}p_1^2A_1^2 + \cdots\right)U\Bigg] \\ \times|\psi_i\rangle\phi(q_1)\phi(q_2).\end{aligned} \tag{10}$$

Putting $p=-i\partial/\partial q$, we obtain

$$\begin{aligned}\Psi_{\mathcal{M}_1\mathcal{M}_2} = F\Bigg(\phi(q_1)\phi(q_2) - g(A_1)_w\phi'(q_1)\phi(q_2) \\ - g(A_2)_w\phi(q_1)\phi'(q_2) + \frac{g^2}{2}(A_1^2)_w\phi''(q_1)\phi(q_2) \\ + \frac{g^2}{2}(A_2^2)_w\phi(q_1)\phi''(q_2) + g^2(A_2,A_1)_w\phi'(q_1)\phi'(q_2) \\ + O(g^3)\Bigg),\end{aligned} \tag{11}$$

where $F=\langle\psi_f|WVU|\psi_i\rangle$, $(A_1)_w=\langle\psi_f|WVA_1U|\psi_i\rangle/F$, $(A_1^2)_w=\langle\psi_f|WVA_1^2U|\psi_i\rangle/F$, $(A_2)_w=\langle\psi_f|WA_2VU|\psi_i\rangle/F$, $(A_2^2)_w=\langle\psi_f|WA_2^2VU|\psi_i\rangle/F$, and $(A_2,A_1)_w$ is defined by

$$(A_2,A_1)_w = \frac{\langle\psi_f|WA_2VA_1U|\psi_i\rangle}{\langle\psi_f|WVU|\psi_i\rangle}. \tag{12}$$

Following measurement of q_1 and q_2, the expected value of their product is given by

$$\langle q_1q_2\rangle = \frac{\int q_1q_2|\Psi_{\mathcal{M}_1\mathcal{M}_2}|^2dq}{\int|\Psi_{\mathcal{M}_1\mathcal{M}_2}|^2dq}. \tag{13}$$

For simplicity, let us make the following assumption (we will discuss the general case later):

Assumption A. The initial pointer distribution ϕ is real valued, and its mean is zero, i.e., $\int q\phi^2(q)dq=0$.

We also assume, without loss of generality, that ϕ is normalized so that $\int\phi^2=1$. With these assumptions, all the terms in Eq. (13) of order 0 and 1 in g vanish, and we are left with

$$\langle q_1 q_2 \rangle = g^2[(A_2,A_1)_w + \overline{(A_2,A_1)}_w + \overline{(A_1)}_w (A_2)_w + (A_1)_w \overline{(A_2)}_w] \times \left(\int q\phi(q)\phi'(q)dq \right)^2, \quad (14)$$

where the overbars denote complex conjugates. Integration by parts implies $\int q\phi(q)\phi'(q)dq = -\frac{1}{2}$, so we obtain the final result

$$\langle q_1 q_2 \rangle = \frac{g^2}{2} \mathrm{Re}[(A_2,A_1)_w + (A_1)_w \overline{(A_2)}_w]. \quad (15)$$

Here $(A_2,A_1)_w$ is the sequential weak value given by Eq. (12); note the reverse order of operators, to fit with the convention of operating on the left.

IV. THE SEQUENTIAL WEAK VALUE

In the section above we considered two measurements—a measurement of A_1 at time t_1 and of A_2 at t_2—and we looked at the product of the outcomes $q_1 q_2$ in the limit when the coupling of the measuring devices with the measured system was weak. This procedure was motivated by our example of the double interferometer: we wanted to check whether the photon followed a given path, say the path that goes along C in the first interferometer and then along E in the second interferometer. In that case the variables of interest are P_C, the projector on path C and P_E, the projector on path E. When the photon follows this path, the value of the *product* of these projectors is 1 while in all other situations the product is 0. We wanted to see what the behavior of the photon was when the measurements did not disturb it significantly.

Since q_1 measures A_1 and q_2 measures A_2, it seems obvious that the quantity that represents the product of the two observables is $\langle q_1 q_2 \rangle$ given in Eq. (14) above. However, the situation is more subtle, as we show below.

Consider the simpler case of two *commuting* operators A_1 and A_2, and suppose we are interested in the value of the product $A_2 A_1$ at some time t. (Note that we are now talking about operators at one given time, not at two different times.) We can measure this product in two different ways. First, we can measure the product directly, by coupling a measuring device directly to the product via the interaction Hamiltonian $H = gpA_2A_1$. When we make the coupling weaker, we find that the pointer indicates the value

$$\langle q \rangle = g\ \mathrm{Re}(A_1 A_2)_w = g\ \mathrm{Re}\frac{\langle \psi_f | A_2 A_1 | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle}. \quad (16)$$

This is straightforward: it is simply the weak value of the operator $A_2 A_1$. On the other hand, we could attempt to measure the product in the same way that we measured the sequential product. That is, we can use two measuring devices with pointer position variables q_1 and q_2, couple the first measuring device to A_1 and the second to A_2, and then look at the product $q_1 q_2$. The latter method was proposed by Resch and Steinberg [7] for the simultaneous measurement of two operators. They showed that in this case

$$\langle q_1 q_2 \rangle = \frac{g^2}{2} \mathrm{Re}[(A_1 A_2)_w + (A_1)_w \overline{(A_2)}_w]. \quad (17)$$

We see that the value indicated by $\langle q_1 q_2 \rangle$ is *not* equal to the weak value of the product, but contains a supplementary term, $\mathrm{Re}(A_1)_w \overline{(A_2)}_w$. In other words, although we expected the two methods to be equivalent, it is not the case. To obtain the true weak value of the product we must subtract this second term. This second term is an artifact of the method of using two separate measuring devices rather than coupling one measuring device directly to the product operator.

In the case of sequential measurement there is no product operator to start with, for we are interested in the product of the values of operators at two different times. Hence the first method, of coupling directly to the product operator, makes no sense, and we must use two independent couplings. In order to obtain the quantity of interest, i.e., the quantity that is relevant to situations such as the double interferometer of Sec. II, we must subtract the term $\mathrm{Re}(A_1)_w \overline{(A_2)}_w$ from Eq. (15). We thus conclude that the quantity of interest is the sequential weak value given in Eq. (12).

V. GENERAL SEQUENTIAL WEAK MEASUREMENT

Sequential weak measurement can be easily extended to n measurements of Hermitian operators A_i with intervening unitary evolution steps U_i. The weak values are given by

$$(A_n, \ldots, A_1)_w = \frac{\langle \psi_f | U_{n+1} A_n U_n \ldots A_1 U_1 | \psi_i \rangle}{\langle \psi_f | U_{n+1} U_n \ldots U_1 | \psi_i \rangle}, \quad (18)$$

and the expected values $\langle q_1 q_2 \ldots q_n \rangle$ can be expressed in terms of these weak values. For example, with assumption A,

$$\langle q_1 q_2 q_3 \rangle = \frac{g^3}{4} \mathrm{Re}[(A_3,A_2,A_1)_w + (A_2,A_1)_w \overline{(A_3)}_w + (A_3,A_1)_w \overline{(A_2)}_w + (A_3,A_2)_w \overline{(A_1)}_w], \quad (19)$$

and the case of general n is given in the Appendix. Similarly, we can express expected values for products of momenta in terms of the weak values (see the Appendix). For instance,

$$\langle p_1 p_2 \rangle = 2(gv)^2\ \mathrm{Re}[-(A_2,A_1)_w + (A_1)_w \overline{(A_2)}_w]. \quad (20)$$

Mixed products of positions and momenta give similar formulas. For instance,

$$\langle q_1 p_2 \rangle = -g^2 v\ \mathrm{Im}[(A_2,A_1)_w + \overline{(A_1)}_w (A_2)_w]. \quad (21)$$

The foregoing examples illustrate a general pattern, which is that expectations of products of p's and q's depend on the real part of sequential weak values if there is an even number of p's in the product and on the imaginary part if there is an odd number of p's.

The sequential weak values satisfy the following rules:

(1) *Linearity in each variable separately.*

$$(A_n, \ldots, A_i, \ldots, A_1)_w + (A_n, \ldots, A_i', \ldots, A_1)_w = (A_n, \ldots, (A_i + A_i'), \ldots, A_1)_w,$$

for any $1 \le i \le n$.

(2) *Agreement with strong measurement.* Suppose that, with preselection by $|\psi_i\rangle$ and postselection by $|\psi_f\rangle$, strong measurements of $A_1, A_2, \ldots, A_n$ always give the same out-

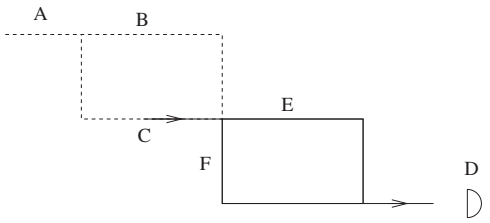

FIG. 3. The double interferometer restricted to its second interferometer. According to Eq. (22), the ratio of the weak values E_w/F_w in the second interferometer, with photons injected along C, is the same as the ratio of the sequential weak values $(E,C)_w/(F,C)_w$ in the double interferometer with photons injected along A.

comes $a_1, a_2, \ldots, a_n$; then $(A_n, \ldots, A_1)_w = a_1 a_2 \cdots a_n$.

(3) *Marginals*. If I is the identity operator at location i,

$$(A_n, \ldots, A_{i+1}, A_{i-1}, \ldots, A_1)_w = (A_n, \ldots, A_{i+1}, I, A_{i-1}, \ldots, A_1)_w.$$

We can illustrate some of these rules with the double interferometer experiment (Fig. 1). The measurements we consider are projectors that detect the presence of a photon on various edges; for instance, the projector P_B indicates whether a photon is present on the edge B. For simplicity we write B_w for the weak value $(P_B)_w$, etc., and we use the same convention for sequential weak values. Then using Eq. (7) we find $C_w=1$, $B_w=0$, $E_w=1$, and $F_w=0$. Using Eq. (12) we find $(E,B)_w=1/2$, $(F,B)_w=-1/2$, $(E,C)_w=1/2$, and $(F,C)_w=1/2$. Since $P_E+P_F=I$, rule (1) implies $(E,B)_w+(F,B)_w=(I,B)_w$, and then rule (3) implies $(I,B)_w=B_w$. Thus we expect $(E,B)_w+(F,B)_w=B_w$, which holds if we substitute the values above. Similarly $(E,C)_w+(F,C)_w=1/2+1/2=C_w$, and so on. As for rule (2), we have seen (Sec. II) that strong measurement of P_C and P_E yields 1, so we expect the weak values to be the same, as is the case.

There is a further rule that applies when one of the operators being measured is a projector. We illustrate it with the double interferometer. We can write

$$\frac{(E,C)_w}{(F,C)_w} = \frac{\langle D|U_3 P_E U_2|C\rangle\langle C|U_1|A\rangle}{\langle D|U_3 P_F U_2|C\rangle\langle C|U_1|A\rangle} = \frac{\langle D|U_3 P_E U_2|C\rangle}{\langle D|U_3 P_F U_2|C\rangle} = \frac{E_w}{F_w}. \tag{22}$$

Here E_w and F_w in the final ratio are calculated assuming that $|\psi_i\rangle=|C\rangle$, in other words, as though we were calculating weak values for the second interferometer treated separately from the rest of the system, with initial state $|C\rangle$ and postselection by $|D\rangle$ (Fig. 3). If we only knew the single-measurement weak values E_w, F_w, and C_w, we could calculate $(E,C)_w$ and $(F,C)_w$ using this rule and the relationship $(E,C)_w+(F,C)_w=C_w$ derived above.

VI. BROADENING THE CONCEPT: WEAK INTERACTIONS

So far, we have considered ideal weak measurements, in which the pointer distribution is real and has zero mean (assumption A). If we drop these assumptions, we find in place of Eq. (2) that

$$\langle q\rangle = \mu + g(\mathrm{Re}[A_w] + \mathrm{Im}[A_w]y), \tag{23}$$

where $y=\int \bar{\phi}(pq+qp)\phi dq - 2\mu\nu$, with $\mu=\int \bar{\phi} q \phi dq$, $\nu=\int \bar{\phi} p \phi dq$.

The expectation $\langle r_1 r_2 \cdots r_n\rangle$ for a general initial pointer distribution, where each r_i is either q_i or p_i, is a very complicated expression, but, so far as the system goes, depends only on the real and complex parts of sequential weak values up to $(A_n, \ldots, A_1)_w$. Thus we can write

$$\langle r_1 r_2 \cdots r_n\rangle = \Phi(\mathrm{Re}(A_n, \ldots, A_1)_w, \mathrm{Im}(A_n, \ldots, A_1)_w, \ldots, \mathrm{Re}(A_n)_w, \mathrm{Im}(A_n)_w, \ldots, \mathrm{Re}(A_1)_w, \mathrm{Im}(A_1)_w), \tag{24}$$

for some polynomial function Φ. The coefficients in Φ are themselves polynomials in expectations $\int \bar{\phi}\gamma(p_i,q_i)\phi dq$ for polynomials γ, as we see in the case of Eq. (23), where y has this form.

In the next section, we shall want to consider the most general possible type of *weak interaction* which allows any sort of (suitably weak) coupling between the system and an ancilla followed by any further evolution or measurement of the ancilla alone (the pointer in our previous discussion and its von Neumann measurement interaction gpA will be a special case of such an ancilla and weak interaction). Our notion of general weak interaction is the following: Consider the system and ancilla initially in product state $|\psi_i\rangle|\xi\rangle$. Let $H_{\mathrm{S,anc}}$ be any Hamiltonian of the joint system, and g a coupling constant. For a single interaction event, and to first order in g, the state becomes

$$(I - igH_{\mathrm{S,anc}})|\psi\rangle|\xi\rangle. \tag{25}$$

Any joint Hamiltonian may be expressed as a sum of products of individual Hamiltonians

$$H_{\mathrm{S,anc}} = \sum_k H_{\mathrm{S}}^k \otimes H_{\mathrm{anc}}^k. \tag{26}$$

Postselecting the system state in Eq. (25) with $|\psi_f\rangle$ gives

$$\Psi_{\mathrm{anc}} = \langle\psi_f|\psi_i\rangle\left(I_{\mathrm{anc}} - ig\sum_k (H_{\mathrm{S}}^k)_w H_{\mathrm{anc}}^k\right)|\xi\rangle. \tag{27}$$

The system Hamiltonians H_{S}^k have been effectively replaced by their weak values $(H_{\mathrm{S}}^k)_w$. The important point here is that all subsequent manipulations of the ancilla will depend on the preselected and postselected system only through *weak values* of suitably chosen observables. A similar result clearly holds for any sequential weak interactions and suitably associated sequential weak values, and also for terms of any higher order in g.

As a simple illustrative example, suppose that the ancilla is the pointer system of a von Neumann measurement interaction with assumption A in force, and that this same pointer is weakly coupled twice for the sequential measurement of both A_1 and A_2. If this pointer has position q and momentum p, the pointer state after postselection is

$$\Psi_{\mathcal{M}} = \langle\psi_f|(U_3 e^{-igpA_2} U_2 e^{-igpA_1} U_1)|\psi_i\rangle\phi(q), \tag{28}$$

yielding

$$\langle q \rangle = g\ \mathrm{Re}[(A_1)_w + (A_2)_w].$$

The effect in this instance is therefore the same as adding the individual postmeasurement results, and it depends on the system only through associated weak values.

VII. COUNTERFACTUALITY AND WEAK MEASUREMENT

Counterfactual computation [8,9] provides a general framework for looking at counterfactual phenomena, including interaction-free measurement as a special case. We consider arbitrary protocols, at various points of which a quantum computer can be inserted. The computer has a switch qubit (with $|0\rangle$=off and $|1\rangle$=on) and an output qubit. A special case of this formalism is where the protocol is represented by an optical circuit, and a computer insertion means that the computer (or a copy of it) is placed in some path of the circuit and is switched on by a photon passing along that path.

We assume that the computer is programmed ready to perform a computational task with answer 0 or 1 which will be written into the output qubit if the switch is turned on. In addition to the switch and output qubits, the protocol will in general have additional qubits, and will involve some measurements. We say that an outcome of these measurements *determines the computer output* if that outcome only occurs when the computer output has a specific value, $|0\rangle$ or $|1\rangle$. Such an outcome is said to be counterfactual if its occurrence also implies that the computer was never switched on, i.e., its switch was never set to $|1\rangle$, during the protocol.

To make this precise, note first that one can always produce an equivalent protocol in which the state is entangled with extra qubits and the measurement deferred to the end of the protocol. Thus the protocol can be assumed to consist of a period of unitary evolution followed by a measurement, which can be assumed (again by adding extra qubits) to be a projective measurement. Let $|\psi_i\rangle$ be the initial state of the protocol, and let $|\psi_f\rangle$ be a measurement outcome that determines some specific computer output, in the sense defined above. Suppose the computer is inserted n times. Let $\mathcal{F}$ (for "off") denote the projection $|0\rangle\langle 0|$ onto the off value of the computer switch and $\mathcal{N}$ (for "on") denote the complementary projector $|1\rangle\langle 1|$, and let ξ be one of the 2^n possible strings of $\mathcal{F}$'s or $\mathcal{N}$'s of length n; we call this a *history*. Let U_i denote the unitary evolution in the protocol between the $(i-1)$th and ith insertions of the computer.

Definition VII.1. (Counterfactuality by histories [9].) The measurement outcome $|\psi_f\rangle$ is a counterfactual outcome if (1) $|\psi_f\rangle$ determines the computer output and (2) the amplitude of any history ξ containing an $\mathcal{N}$ vanishes. In other words, for all histories ξ other than the all-$\mathcal{F}$ history, $\langle\psi_f|U_{n+1}\xi_n U_n \ldots U_2\xi_1 U_1|\psi_i\rangle=0$.

One may question whether this is the "correct" definition of a notion of counterfactual computation or whether alternative definitions might be convincingly plausible. Condition (1) is uncontroversial but condition (2) might seem less immediately compelling. It is evidently equivalent to obtaining a null result if we carry out a strong nondemolition measurement of $\mathcal{N}$ at each computer insertion. However the disturbance that such a measurement causes might lead one to question the suitability of this condition. Indeed recently Hosten *et al.* [10] proposed an alternative definition of counterfactual computation that violates condition (2) of definition VII.1 and sparked a controversy [11] over the relative merits and validity of the two notions. We will now develop some alternative characterizations of our definition VII.1 in terms of *weak* measurements, thereby addressing the disturbance issue. We will argue that these characterizations considerably strengthen the credibility of the original definition as the "correct" one.

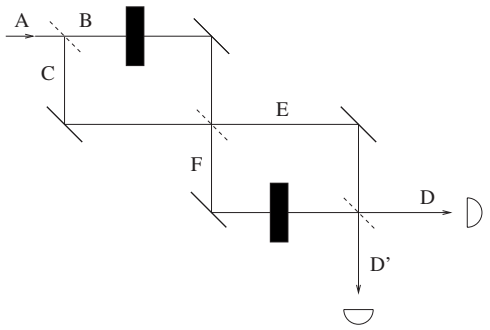

FIG. 4. The double interferometer of Fig. 1 treated as a protocol with computer insertions (black rectangles) in paths B and F. If a photon passes down either of these paths, the computer runs.

Let us therefore consider carrying out a weak measurement of $\mathcal{N}$ at each insertion. A nonzero weak value implies that there is a detectable physical effect that can only occur if the computer is switched on. Vaidman's treatment of the three-box paradox [12] gives a good example of this reasoning.

Our two-interferometer example shows that it does not suffice to consider the individual weak values at each insertion. Suppose the computer is inserted in paths B and F, as shown in Fig. 4. Then we have seen that the weak values B_w and F_w are zero, yet the sequential weak value $(F,B)_w$ is nonzero. The nonvanishing of the sequential weak value implies that a photon passes along *both* path B and F, since there is a physical effect that causes correlated deflections of pointers at both sites.

There is a subtlety here, because it could be argued that because sequential pairwise weak measurements give second-order effects in g [see Eq. (15)], we might detect a departure from zero in the weak measurements for each operator individually, i.e., in the deflections of the pointers at B and F, if we looked at second- or higher-order terms in g. However, if A is any projector and $A_w=0$, then the von Neumann interaction e^{-igpA} reduces to $Ae^{-igp}+I-A$, which is the identity to all orders in g in the weak measurement calculation. Thus we truly need to carry out the sequential weak measurement here to identify the physical effect due to the photon.

In general, we need to consider all possible sequential weak measurements to obtain an adequate test of counterfactuality.

We therefore propose the following:

Definition VII.2. (Counterfactuality by weak values.) The

measurement outcome $|\psi_f\rangle$ is a counterfactual outcome if (1) $|\psi_f\rangle$ determines the computer output and (2) $(\mathcal{N}_{i_k},\mathcal{N}_{i_{k-1}},\ldots\mathcal{N}_{i_1})_w=0$, for any $1\leq i_1<i_2<\cdots<i_k\leq n$, where n is the number of insertions of the computer.

By Eq. (18), conditions (2) for VII.1 and VII.2 are equivalent, using the fact that $\mathcal{F}+\mathcal{N}=1$ together with the linearity and marginal rules. For instance, with two insertions of the computer, condition (2) of definition VII.1 amounts to $(\mathcal{N}_1,\mathcal{N}_2)_w=0$, $(\mathcal{F}_1,\mathcal{N}_2)_w=0$, and $(\mathcal{N}_1,\mathcal{F}_2)_w=0$, and these imply $(\mathcal{N}_1)_w=0$, $(\mathcal{N}_2)_w=0$, and $(\mathcal{N}_1,\mathcal{N}_2)_w=0$, which constitute condition (2) for definition VII.2.

We can try to strengthen the requirements for counterfactuality by demanding that a zero response is obtained for any conceivable weak interaction, in the sense of the preceding section. In our present application we must further restrict the weak interaction to take place only if the switch has the property of being "on," i.e., the interaction Hamiltonian must have the form $(\mathcal{N}\otimes I_{\text{anc}})H_{\text{s,anc}}(\mathcal{N}\otimes I_{\text{anc}})$. We say that such an interaction is a weak interaction involving the projector $\mathcal{N}$. Since $\mathcal{N}$ is a one-dimensional projector, this implies that the interaction Hamiltonian has the form $\mathcal{N}\otimes H_{\text{anc}}$. In a more general scenario the projector $\tilde{\mathcal{N}}$ for counterfactuality (analogous to the switch being "on") may have rank larger than 1 and then the interaction Hamiltonian may have the more general form $(\tilde{\mathcal{N}}\otimes I_{\text{anc}})M_{\text{s,anc}}(\tilde{\mathcal{N}}\otimes I_{\text{anc}})$ for any Hermitian M. For example, the switch may be a photon with both path and polarization properties. Then a weak interaction restricted to its presence on a path would correspond to a two-dimensional projector on its polarization state space associated to that path.

Definition VII.3. (Counterfactuality by general weak interactions.) The measurement outcome $|\psi_f\rangle$ is a counterfactual outcome if (1) $|\psi_f\rangle$ determines the computer output and (2) Any possible weak interaction involving the projections $\mathcal{N}_1,\ldots,\mathcal{N}_n$ yields a null result.

By a null result, we mean the same result that would be obtained for $g=0$. It is not difficult to show that this apparently much broader concept is in fact equivalent to definition VII.2. In one direction, we know from the last section that any expectation depends only on the sequential weak values, involving the projectors $\mathcal{N}_i$, so when these weak values vanish we obtain a null result. In the other direction, we have only to show that we can choose particular weak interactions whose null results will imply the vanishing of all sequential weak values. However, if we first obtain a null value of $\langle q_i\rangle$ and $\langle p_i\rangle$ for the standard von Neumann measurement weak interaction for every i, then we know by Eqs. (2) and (3) that both real and imaginary parts of all the weak values $(\mathcal{N}_i)_w$ are zero. Then by obtaining null values of $\langle q_iq_j\rangle$ and $\langle p_iq_j\rangle$ for all $i<j$, we infer from Eqs. (15) and (21) that the real and imaginary parts of all $(\mathcal{N}_j,\mathcal{N}_i)_w$ are zero. We continue this way, using the fact that expectations of products of p's and q's with an even number of p's depend on the real part of sequential weak values, whereas those with an odd number of p's depend on their imaginary parts (see the Appendix).

We have therefore proved:

Theorem VII.4. All three definitions, VII.1, VII.2, and VII.3, are equivalent.

VIII. DISCUSSION

The lesson that we learn from our results is that there is a very interesting structure in quantum mechanics. When we perform measurements to find out which way photons go through the double interferometer, and when we make these measurement weak enough, then (given that the final postselection is successful) the results we obtain are consistent yet very strange. Indeed the measurements indicate that $N/2$ photons go along path CF but also that $-N/2$ photons go along path BF; strange as this number $-N/2$ is, it nonetheless combines with the result for the path $N/2$ to imply, correctly, that no photons pass along path F. This consistency applies to all the measurement results we obtain, and is very reminiscent of the pattern of weak measurements seen in Hardy's paradox [1], where negative numbers of particles are also obtained (though here simultaneous [7] rather than sequential weak measurements are carried out).

What are we to make of these strange yet consistent results? The bold assumption is that, as long as they truly give consistent answers in every physical situation, then they are the actual values of the parameters being measured. And in fact there is a body of work showing that weak values, even when they lie in an unexpected range, can be treated as though they were the actual values in the underlying physical theory, and that they yield correct predictions. Examples include weakly measured negative kinetic energies when a particle is in a classically forbidden region [13], and weakly measured faster-than-light velocities that are associated with Cerenkov radiation [14]. Here we are looking at traditional weak values at a single time, as in Eq. (1). For sequential weak values, we can make a similar argument. The double interferometer already gives an example that illustrates their consistency. We thus suggest that sequential weak values should also be interpreted as truly representing actual values of the parameters being measured, providing valuable insights in further physical situations.

To conclude, we mention the striking fact that sequential weak values are formally closely related to amplitudes. Consider the case where we measure n projectors $P_{X_1},\ldots,P_{X_n}$ that define a path π_x between the initial and postselected states $|\psi_i\rangle$ and $|\psi_f\rangle$, respectively. We can write

$$(P_{X_n},\ldots,P_{X_1})_w=\frac{\langle\psi_f|U_{n+1}|X_n\rangle\langle X_n|U_n|X_{n-1}\rangle\cdots\langle X_1|U_1|\psi_i\rangle}{\langle\psi_f|U_n\cdots U_1|\psi_i\rangle}=\frac{\text{amplitude}(\pi_x)}{\sum_i \text{amplitude}(\pi_i)},\tag{29}$$

where π_i runs over all paths between $|\psi_i\rangle$ and $|\psi_f\rangle$. Nonetheless, weak values are similar to measurement results rather

than amplitudes. This way of looking at sequential weak values suggests a close connection with path integrals that remains to be explored.

ACKNOWLEDGMENTS

We thank L. Vaidman for helpful comments on an earlier version. G.M. acknowledges support from the project PROSECCO (Contract No. IST-2001-39227) of the IST-FET programme of the EC. R.J. and S.P. are supported by the EPSRC QIPIRC and EC networks QAP and QICS, and S.P. also acknowledges support from EPSRC Grant No. GR/527405/01.

APPENDIX: CALCULATION OF GENERAL CORRELATIONS

With assumption A, we show here that the general version of Eq. (15) is

$$\langle q_1 q_2 \cdots q_n \rangle = \frac{g^n}{2^{n-1}} \mathrm{Re} \sum_{r \geq s} \sum_{\mathbf{i},\mathbf{j}} (A_{i_r}, \ldots, A_{i_1})_w \overline{(A_{j_s}, \ldots, A_{j_1})_w}, \tag{A1}$$

where the weak values in this formula are given by Eq. (18). In Eq. (A1) the sum is over all ordered indices $\mathbf{i}=(i_1,\ldots,i_r)$ with $i_p < i_{p+1}$ for $1 \leq p \leq r-1$, and ordered indices $\mathbf{j}=(j_1,\ldots,j_s)$ that make up the complement of $\mathbf{i}$ in the set of integers from 1 to n, i.e., that satisfy $(i_1,\ldots,i_r) \cup (j_1,\ldots,j_s)=(1,2,\ldots,n)$ and $(i_1,\ldots,i_r) \cap (j_1,\ldots,j_s)=\emptyset$. We include the empty set $\emptyset$ as a possible set of indices. In order not to count indices twice, we require $r \geq s$, and when $r=s$ we require $i_1=1$.

For instance, with $n=2$, the possible indices are $\mathbf{i}=(1,2)$, $\mathbf{j}=\emptyset$; $\mathbf{i}=(1)$, $\mathbf{j}=(2)$, which yields

$$\langle q_1 q_2 \rangle = \frac{g^2}{2} \mathrm{Re}[(A_2,A_1)_w + (A_1)_w \overline{(A_2)_w}]. \tag{A2}$$

This is just Eq. (15). For $n=3$ we have $\mathbf{i}=(1,2,3)$, $\mathbf{j}=\emptyset$; $\mathbf{i}=(1,2)$, $\mathbf{j}=(3)$; $\mathbf{i}=(1,3)$, $\mathbf{j}=(2)$; $\mathbf{i}=(2,3)$, $\mathbf{j}=(1)$, giving Eq. (19). Equation (A1) is proved in the same way as Eq. (15), the state of the n pointers after postselection being

$$\begin{aligned}
\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n} &= \langle \psi_f | (U_{n+1} e^{-igp_n A_n} U_n \cdots U_2 e^{-igp_1 A_1} U_1) \\
&\quad \times |\psi_i\rangle \phi(q_1) \cdots \phi(q_n) \\
&= \langle \psi_f | \{U_{n+1}[\phi(q_n) - gA_n \phi'(q_n) + \cdots] U_n \cdots U_2 \\
&\quad \times [\phi(q_1) - gA_1 \phi'(q_1) + \cdots] U_1\} | \psi_i \rangle \\
&= \langle \psi_f | U_{n+1} U_n \cdots U_1 | \psi_i \rangle \Bigg(1 + g \sum_i \frac{\phi'(q_i)}{\phi(q_i)} (A_i)_w \\
&\quad + g^2 \sum_{i<j} \frac{\phi'(q_i)\phi'(q_j)}{\phi(q_i)\phi(q_j)} (A_j,A_i)_w + \cdots \Bigg) \\
&\quad \times \phi(q_1) \cdots \phi(q_n).
\end{aligned} \tag{A3}$$

Assumption A implies that only the terms in $q_1 q_2 \cdots q_n$ in $|\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n}|^2$ need to be taken into account in calculating

$$\langle q_1 q_2 \cdots q_n \rangle = \frac{\int q_1 q_2 \cdots q_n |\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n}|^2 dq_1 \cdots dq_n}{\int |\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n}|^2 dq_1 \cdots dq_n},$$

and this leads to Eq. (A1).

We can also calculate $\langle p_1 p_2 \cdots p_n \rangle$, the product of the momenta of the pointers. To do this, it is convenient to move to the momentum basis, replacing $\phi(q)$ by its Fourier transform $\tilde{\phi}(p)$ and carrying out an expansion in the p_i,

$$\begin{aligned}
\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n} &= \langle \psi_f | (U_{n+1} e^{-igp_n A_n} U_n \cdots U_2 e^{-igp_1 A_1} U_1) \\
&\quad \times |\psi_i\rangle \tilde{\phi}(p_1) \cdots \tilde{\phi}(p_n) \\
&= \langle \psi_f | U_{n+1} U_n \cdots U_1 | \psi_i \rangle \Bigg(1 - ig \sum_i p_i (A_i)_w \\
&\quad + (-ig)^2 \sum_{i<j} p_i p_j (A_j,A_i)_w + \cdots \Bigg) \tilde{\phi}(p_1) \cdots \tilde{\phi}(p_n).
\end{aligned} \tag{A4}$$

Assumption A implies that only the terms in $p_1 p_2 \cdots p_n$ in $|\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n}|^2$ need be considered in calculating

$$\langle p_1 p_2 \cdots p_n \rangle = \frac{\int \overline{\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n}} p_1 \cdots p_n \Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n} dp_1 \cdots dp_n}{\int |\Psi_{\mathcal{M}_1 \cdots \mathcal{M}_n}|^2 dp_1 \cdots dp_n}. \tag{A5}$$

It is simplest to treat the cases of n even and odd separately. For the even case we have

$$\begin{aligned}
\langle p_1 p_2 \cdots p_{2m} \rangle &= 2(-1)^m (gv)^{2m} \mathrm{Re} \sum_{r \geq s} \sum_{\mathbf{i},\mathbf{j}} (-1)^r \\
&\quad \times (A_{i_r}, \ldots, A_{i_1})_w \overline{(A_{j_s}, \ldots, A_{j_1})_w}
\end{aligned} \tag{A6}$$

and for the odd case

$$\begin{aligned}
\langle p_1 p_2 \cdots p_{2m+1} \rangle &= 2(-1)^{m+1} (gv)^{2m+1} \mathrm{Im} \sum_{r>s} \sum_{\mathbf{i},\mathbf{j}} (-1)^r \\
&\quad \times (A_{i_r}, \ldots, A_{i_1})_w \overline{(A_{j_s}, \ldots, A_{j_1})_w},
\end{aligned} \tag{A7}$$

where $v = \int p^2 \tilde{\phi}^2(p) dp$.

The case of mixed products of positions and momenta are treated similarly, and they depend only on the real or imaginary parts of the sequential weak values given by Eq. (18). For example, to calculate $\langle q_1 p_2 \rangle$ we express the first variable in the position basis and the second in the momentum basis,

$$\Psi_{\mathcal{M}_1,\mathcal{M}_2} = \langle \psi_f | U_3 U_2 U_1 | \psi_i \rangle [\phi(q_1)\tilde{\phi}(p_2) + g(A_1)_w \phi'(q_1)\tilde{\phi}(p_2) - ig(A_2)_w \phi(q_1) p_2 \tilde{\phi}(p_2) + ig^2 (A_2, A_1)_w \phi'(q_1) p_2 \tilde{\phi}(p_2)],$$

which yields Eq. (21). For these mixed products, since there is a factor of i for each p in the product, we take the imaginary part of weak values when there is an odd number of p's present and the real part otherwise.

Thus all possible expectations of products of position or momentum can be obtained from the sequential weak values.

[1] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, in *Proceedings of NATO ARW Mykonos 2000 Decoherence and its Implications in Quantum Computation and Information Transfer*, NATO Advanced Research Workshop, Mykonos, Greece, 2000, edited by A. Gonis and P. Turchi (IOS Press, Amsterdam, The Netherlands, 1991).

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

[3] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes* (Wiley-VCH, Weinheim, Germany, 2005).

[4] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).

[5] B. Bläsi and L. Hardy, Phys. Lett. A **207**, 119 (1995).

[6] A. C. Elitzur and L. Vaidman, Found. Phys. **23**, 987 (1993).

[7] K. J. Resch and A. M. Steinberg, Phys. Rev. Lett. **92**, 130402 (2004).

[8] R. Jozsa, Chaos, Solitons Fractals **10**, 1657 (1999).

[9] G. Mitchison and R. Jozsa, Proc. R. Soc. London, Ser. A **457**, 1175 (2001).

[10] O. Hosten, M. T. Rakher, J. T. Barreiro, N. A. Peters, and P. G. Kwiat, Nature (London) **439**, 949 (2006).

[11] G. Mitchison and R. Jozsa, e-print arXiv:quant-ph/0606092.

[12] L. Vaidman, Phys. Rev. Lett. **98**, 160403 (2007).

[13] D. Rohrlich, Y. Aharonov, S. Popescu, and L. Vaidman, Ann. N.Y. Acad. Sci. **755**, 394 (1995).

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