What Is a Quantum-Mechanical "Weak Value" the Value of?

Bengt E.Y. Svensson

Abstract A so called "weak value" of an observable in quantum mechanics (QM) may be obtained in a weak measurement + post-selection procedure on the QM system under study. Applied to number operators, it has been invoked in revisiting some QM paradoxes (*e.g.*, the so called Three-Box Paradox and Hardy's Paradox). This requires the weak value to be interpreted as a *bona fide* property of the system considered, *a par* with entities like operator mean values and eigenvalues. I question such an interpretation; it has no support in the basic axioms of quantum mechanics and it leads to unreasonable results in concrete situations.

1 Introduction

The pioneering work by Aharonov, Albert and Vaidman [1] introduced the concept of 'weak measurements' of an observable in quantum mechanics (QM) in combination with the procedure of 'post-selection'. The result of such a weak measurement + post-selection can be expressed in terms of the 'weak value' $(_f \hat{O}_{in})_w$ of the observable O under study. For an initial ('pre-selected') state $|in\rangle$ and a final ('post-selected') state $|f\rangle$, this weak value is defined by[1]

---

[1]It can be measured in a so called indirect measurement scheme in which the examined 'system' starts in the initial state $|in\rangle$ ('pre-selection'), interacts weakly with a suitable 'meter', and is then projected into a final state $|f\rangle$ ('post-selection'). Following this procedure, the weak value can then be obtained by suitably reading off the meter. See, *e.g.*, [2–4] for further details.

B.E.Y. Svensson (✉)
Theoretical High Energy Physics, Department of Astronomy and Theoretical Physics, Lund University, Sölvegatan 14A, 223 62 Lund, Sweden
e-mail: Bengt_E_Y.Svensson@thep.lu.se

$$(_f \hat{O}_{in})_w := \frac{\langle f | \hat{O} | in \rangle}{\langle f | in \rangle} \qquad (1)$$

The basic idea of weak measurement + post-selection has since attracted much interest; for recent reviews with further references see, *e.g.*, [2–4]. Not the least has the field opened up new tools for experimentalists to investigate aspects of phenomena that were thought impossible earlier. These include determining (even if only statistically) the trajectories in a double slit experiment without destroying the interference pattern [5], and directly measuring the wave function [6]. The use of the technique for amplification purposes has been spectacular [7–9]. Weak measurements have also been employed to illuminate the fundamental difference between classical and quantum mechanics exhibited by violation of Leggett-Garg inequalities [3, 10–14], also called "Bell inequalities in time".

In addition, weak values of number operators have been invoked to revisit what are conceived as QM paradoxes, like the so called Three-Box Paradox [15–18] and Hardy's Paradox [19–22]. It is even claimed that some of the so called 'counterfactual' statements in these 'paradoxes' can be made 'factual' by using weak measurements, quantified in terms of weak values, instead of the ordinary 'strong' measurements. These claims rely on an straight-forward interpretation of a weak value on *a par* with, *e.g.*, a usual mean value. That is, one ascribes a meaning to a weak value as an 'ordinary' property—like a probability, the number of particles, *etc.*—of the system under investigation.

In this paper, I investigate these very basic interpretational issues. In particular, I criticize such a realistic, straightforward interpretation (for short "RSFI" in the sequel). My emphasis is on a very fundamental (and also elementary) level: given that a weak value $(_f \hat{O}_{in})_w$ has the definition and the operational connotations that it has, are there grounds for seeing it as a *bona fide*, 'ordinary' property (like a mean value $\langle \hat{O} \rangle$; see Sect. 2 for a more explicit description of what I mean by an 'ordinary' property)? I approach this problem from different angles: Do the conventional, basic rules—the 'axioms'—of QM have any bearing? Is an RSFI consistent with varying the entities defining $(_f \hat{O}_{in})_w$, in particular with varying the post-selected state $|f\rangle$?

Note that I do not here question the weak value as a measureable entity pertaining to the system under study, nor the many ingenious ways experimentalist have exploited the weak measurement + post-selection idea ([5–9, 11–14]; see also [3] for a brief review). What I do question is whether a weak value can be given an 'ordinary' meaning.

My focus is on the use of weak values for number operators. However, I begin in Sect. 2 with some more general observations and remarks on weak values. In Sects. 3–5, I then study how weak values of number operators are used in connection with (variants of) the Three-Box Paradox, Hardy's Paradox and a Mach-Zehnder interferometer, and show that an RSFI of these weak values leads to results which seem to lack a reasonable meaning. An appendix is devoted to the relation between weak values of number operators and the so called Aharonov-Bergmann-Lebowitz (ABL) probability formula [23].

2 Interpreting a Weak Value—Some General Remarks

In the application of weak values, it is often taken for granted, explicitly or implicitly, that en entity like $({}_f \hat{O}_{in})_w$ has the same basic meaning as an average value $\langle \hat{O} \rangle$. This is done by Aharonov and collaborators when they treat the Three-Box Paradox and the Hardy's Paradox as I will describe in some detail in Sects. 3 and 4 below. It is also what they do when they discuss "negative kinetic energy" in tunneling experiments [15]. And Vaidman [24, 25] has argued that this is legitimate:

> *If we are certain that a procedure for measuring a certain variable will lead to a definite shift of the unchanged probability distribution of the pointer, then there is an element of reality: the variable equal to this shift.* [24]

Sometimes they or other protagonist of the weak-measurement + post-selection procedure even state that there are experimental support for ascribing such an 'ordinary' meaning to the weak value [2, 25–28].

Ever since the path-breaking paper [1] by Aharonov *et al.* there has been much discussion of what a 'weak measurement' really can accomplish and what significance one can and should ascribe to a 'weak value'; the whole discussion can be retrieved from [2, 15] and references therein. Of the previous criticisms of weak values that seems most similar to mine, I should mention [29, 30] (to which Aharonov and Vaidman replied [31]) and [32]. Of the more recent ones, I emphasize [33] and, even more [34], in which Kastner explicitly points out that a weak value "should be thought of as an amplitude and not an expectation value at all." Indeed, the weak value is the ratio of two amplitudes. My arguments may be seen as a further deepening of Kastner's criticism.

It is without question that the weak value $({}_f \hat{O}_{in})_w$ is an "element of reality" in the general sense described by Vaidman [24] in the quotation above: the weak value is an experimentally accessible entity. And part of the ingenuity of introducing the weak measurement + post-selection procedure is that it allows direct measurement of such a ratio of two amplitudes. But the very fact that an entity can be measured does in no way entails its interpretation. Experiments really have little bearing on the *interpretation* of the weak value; it must be the theory that decides what *meaning* to ascribe to it. Nor can the fact that experiments confirm some of the predictions for weak values be taken as an argument for the meaning of the weak value. For examples, experiments on Hardy's Paradox [21, 22] are indeed ingenious in their own right. However, what they find is but a consequence of QM: they do confirm the results of Aharonov *et al.*, results that are entirely based on QM. To put it in another way: should the experimentalists have found disagreement with the results derived by Aharonov *et al.*, they would have found a violation of QM.

In conclusion so far, the facts that $({}_f \hat{O}_{in})_w$ can be measured and that the results in concrete experimental situations agree with theoretical predictions cannot be taken as support for an RSFI of $({}_f \hat{O}_{in})_w$ as an 'ordinary' property.

In searching for an argument that could motive what meaning to ascribe to $({}_f \hat{O}_{in})_w$, being the ratio of two amplitudes, it is natural to go back to the very foundations of QM, to the basic rules that govern all the endeavor: QM (for a finite system) lives in a (separable) Hilbert space $\mathcal{H}$, an observable O is represented by a self-adjoint

operator $\hat{O}$ in $\mathcal{H}$, the result of (strongly) measuring the observable O is an eigenvalue o_n of $\hat{O}$, a state is describes by a density matrix $\hat{\rho}$ in $\mathcal{H}$, and the probability of obtaining a particular eigenvalue o_n equals the trace $\mathrm{Tr}(\hat{P}_n\hat{\rho})$ with $\hat{P}_n$ the projector onto the subspace of $\mathcal{H}$ spanned by the eigenvectors of $\hat{O}$ with eigenvalue o_n. These rules imply, *i.e.*, that one may ascribe experimental and conceptual meaning to theoretical entities like the eigenvalues of $\hat{O}$, the probability of obtaining an eigenvalue, and the mean value of the operator in a given state. These are examples of what I call the 'ordinary' meaning of the theoretical concepts of QM.

> *Nowhere in these basic rules can I find any motivation for interpreting a ratio of two amplitudes, like the weak value $(_f\hat{O}_{in})_w$, as an 'ordinary' property in this sense.*

But the basic rules are not sacred.[2] No-one would object if they were supplemented with further insight provided this insight were in agreement with all previous experience and with other 'reasonable' arguments, in particular also being compatible with the usual meaning of the concepts involved. To try to find out whether such an approach to the meaning of the weak value $(_f\hat{O}_{in})_w$ is a practical way forward, I will devote the following sections to investigating some concrete applications of weak values to the cases in which the operator $\hat{O}$ is a number operator.

3 The Three-Box Paradox

One (thought) experiment which Aharonov *et al.* analyzed in terms of weak values is the Three-Box Paradox. Imagine, they say [15–17], a single QM particle in any one of three boxes A, B and C. The QM states representing the particle in box A is denoted $|A\rangle$, with corresponding notations for B and C. Let the particle be described by a superposition given by the initial ('pre-selected') state $|in\rangle = (|A\rangle + |B\rangle + |C\rangle)/\sqrt{3}$. Suppose further that the particle is later found in the ('post-selected') state $|f\rangle = (|A\rangle + |B\rangle - |C\rangle)/\sqrt{3}$. Moreover, consider an (intermediate in time) measurement of the projection operator $\hat{A} := |A\rangle\langle A|$. It is a number operator which, upon measurement, counts the number of particles in the box A; its mean value gives the probability of finding the particle in box A. In the same way, $\hat{B} := |B\rangle\langle B|$ is the number operator for the particle in box B, and similarly for C. One is interested in the probability $prob_A$ (in A) for finding the particle in box A when measuring $\hat{A}$, as well as the corresponding probabilities for B and C.

Consider first an ordinary, 'strong' measurement of the respective projection/number operators. The ABL rule as presented in the Appendix applies and gives, for the pre- and post-selected state as given above,

$$prob_A(\text{in } A) = 1 = prob_B(\text{in } B) \quad \text{while} \quad prob_C(\text{in } C) = 1/5 \qquad (2)$$

---

[2]For example, some authors [35, 36] have proposed to consider (essentially) the weak value as an extended probability, which might even take complex values. I shall not follow that track here; it is not what I call "compatible with the usual meaning of the concepts involved".

At first, there seems to be a paradox here. Not only does the total probability to find the particle in any box exceed 1, it is with certainty or at least with probability 1 to be found both in box A and in box B. But the paradox disappears when one realizes that the results apply to *different* projective measurements, which certainly cannot be performed in conjunction without each measurement heavily disturbing ('collapsing') the system and thereby creating totally new conditions for a subsequent measurement. In other words, the paradox only appears if one allows 'counter-factual' statements which require, for their verification, measurements that are non-implementable in an ordinary, 'strong' measurement scheme.

Could weak measurements come to the rescue? It is certainly possible to implement them. Indeed, so the advocates of the weak measurement plus post-selection procedure argue, nothing forbids one to do all three weak measurements of the number operators $\hat{A}$, $\hat{B}$ and $\hat{C}$ successively on the given pre- and post-selected states: the fact that they are weak measurements, so the argument goes, ensures that one may (approximately) disregard measurement disturbances. The conclusion is that, with weak measurements, there are not necessarily any counter-factual statements involved.

The relevant weak values are

$$(_f\hat{A}_{in})_w = 1 = (_f\hat{B}_{in})_w. \tag{3}$$

On the other hand,

$$(_f\hat{A}_{in})_w + (_f\hat{B}_{in})_w + (_f\hat{C}_{in})_w = \left[_f(\hat{A} + \hat{B} + \hat{C})_{in}\right]_w = 1, \tag{4}$$

which together imply

$$(_f\hat{C}_{in})_w = -1! \tag{5}$$

(Of course, this could also be calculated directly from the explicit expression for $(_f\hat{C}_{in})_w$.) Consequently, with an 'ordinary' meaning of these weak values in which one interprets the weak value as a *bona fide* value of a number operator, one arrives at the mind-boggling result that there is minus one particle in box C!

Let me see how this stands further scrutiny.

Let me first note that the result for the weak values of $\hat{A}$ and $\hat{B}$ being unity is in agreement with the strong, ABL-values in Eq. (2) being unity (see the Appendix below).

The 'strong' values of a projection operator are its eigenvalues, 1 and 0. It is this fact that, from the basic rules of QM reviewed in Sect. 2 above, legitimates the result of a measurement of a projection operator to be interpreted as the (relative) number of particles in the respective box, and its mean value as the probability of finding the particle in that box. But how legitimate is it to interpret the *weak value* of a projection operator here as 'the', or even 'a', particle number, or as a probability?

An important difference to an ordinary mean value is that a weak value depends not only on the pre-selected state, $|in\rangle$, but also on the post-selected one, $|f\rangle$. In other words, whatever it is, it is not only referring to the system in the initial, pre-selected state, but to the whole set-up of the situation being analyzed. And, by considering

different combinations of the basis states $|A\rangle$, $|B\rangle$ and $|C\rangle$ for $|f\rangle$, one may get essentially any result for the weak value of a number operator. In fact, one may look upon the post-selected state as a kind of filter which can be tuned as one wants. Expressed differently, even for a given pre-selected state $|in\rangle$, like the one chosen here, there is still the freedom of choosing the post-selected state $|f\rangle$ in any way one likes. The particular one $|f\rangle = (|A\rangle + |B\rangle - |C\rangle)/\sqrt{3}$ used in the formulation above has no preordained physical meaning. It is just one among a multiple continuum of choices, none of which seems more natural than the other. I devote a section in my review [3] to a more extensive study of this ambiguity.

Attempts have also been made by Aharonov and collaborators to give, in terms of thought experiments, physical meaning to, *e.g.*, the value -1 for the weak value of a number operator. However, as I show in [3] for the particular case treated by Rohrlich and Aharonov in [37], their explanation, too, is heavily dependent on the filtering function of post-selection. Therefore, such a thought experiment carries no further explanatory power.

In conclusion, in this case of the so called Three-Box Paradox, there are strong arguments disfavoring an RSFI of the weak values as 'ordinary' values of number operators. Besides not having any motivation from the basic postulates of QM, the weak values here can take essentially any complex value, not in any obvious way related to a number. More conceptually, for a given initial state, the weak values depend also on the choice of the final, post-selected state and cannot be interpreted merely as a property of the initial state alone: a weak value like $(_f\hat{C}_{in})_w$ cannot simply tell how many particles there are in box C.
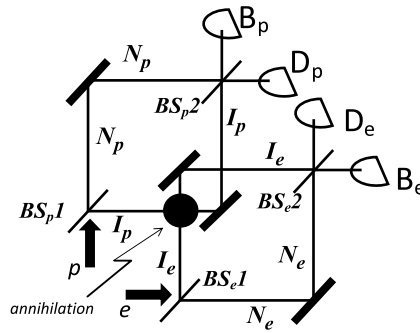
4 Hardy's Paradox

In [19], Hardy outlined a QM experiment that implies a paradox if interpreted in classical terms. Aharonov and collaborators gave an analysis of the experiment [20] in terms of weak values, in which they again relied on an RSFI of weak values of number operators. In this section, I discuss the legitimacy of this interpretation. My treatment here relies heavily on the paper [20] by Aharonov *et al.*

Two experimental groups have performed experiments [21, 22] and confirmed the results of Aharonov *et al.* As I made clear in Sect. 2 above, this has no bearing on the *interpretation* of the weak values.

The setting for Hardy's Paradox is schematically presented in Fig. 1. It is assumed that the beam-splitters are ideal 50–50 splitters, that the arms of the Mach-Zehnder interferometers for the electron respectively for the positron are of equal length, and that there are no other obstructions in the arms but the annihilation region in the interacting arms I_p and I_e where annihilation is assumed to occur with unit probability. Also, the electron and positron are assumed to enter the apparatus simultaneously. The assumption of simultaneous passage of a positron-electron pair through the whole set-up should therefore be valid.

The analysis proceeds by successively considering what happens at the relevant moment of the particles' traversal through the apparatus. In a hopefully self-

Fig. 1 Schematic illustration of the experimental set-up for Hardy's Paradox. An electron (e) and a positron (p) enter each its own Mach-Zehnder interferometer with beam-splitters (BS), and are detected in the B (for 'bright') or the D (for 'dark') ports. They are free to move in the non-interacting arms (N) but annihilate each other in the intersection of the interaction arms (I). The paradox is that a pair appears in the D-ports, indicating that the particles went through the I-arms, even though they should then have been annihilated

explanatory notation, the transition at the beam-splitter $BS_p 1$ is described by

$$|p\rangle \xrightarrow{BS_p 1} = \{|N_p\rangle + i|I_p\rangle\}/\sqrt{2}, \tag{6}$$

with similar transitions at the other beam-splitters. The state of the electron-positron pair just before it enters the second pair of beam-splitters, *i.e.* after the possible annihilation in the I_e-I_p-intersection, is without an $|I_p\rangle \otimes |I_e\rangle$-term, and reads

$$|in\rangle := \{N_p\rangle \otimes |N_e\rangle + i|I_p\rangle \otimes |N_e\rangle + |N_p\rangle \otimes i|I_e\rangle\}/\sqrt{3}. \tag{7}$$

At the second pair of beam-splitters this turns into

$$\{-|D_p\rangle \otimes |D_e\rangle + i|D_p\rangle \otimes |B_e\rangle + i|B_p\rangle \otimes |D_e\rangle - 3|B_p\rangle \otimes |B_e\rangle\}/\sqrt{12}. \tag{8}$$

The paradox is now the following. Noting that the state (8) can be written

$$\{(-|D_p\rangle + i|B_p\rangle) \otimes |D_e\rangle + i|D_p\rangle \otimes |B_e\rangle - 3|B_p\rangle \otimes |B_e\rangle\}/\sqrt{12}$$

$$= \{-i\sqrt{2}|I_p\rangle \otimes |D_e\rangle + i|D_p\rangle \otimes |B_e\rangle - 3|B_p\rangle \otimes |B_e\rangle\}/\sqrt{12}, \tag{9}$$

one sees that a click in the detector arm D_e means that the positron must have traveled through the arm I_p, and similarly for D_p with respect to I_e. Therefore, from a simultaneous click in D_e and D_p it seems reasonable to conclude that both particles should have gone through their respective interacting arm I_p and I_e, in which case they should have annihilated and therefore not have been able to reach any detector. This is Hardy's paradox.

Aharonov *et al.* [20] now want to check this by actually measuring through which arms the particles went. Of course, this cannot be done by an ordinary, 'strong' measurement, since that would 'collapse' the wave-function and thus destroy coherence.

Instead, they invoke weak measurements of number operators, *i.e.*, single particle occupation operators like $|N_p\rangle\langle N_p| =: \hat{N}_p$ and pair occupation operators like $\hat{N}_p \otimes \hat{I}_e$, *etc.* They express their results in terms of the weak values of these operators. The pre-selected state they choose is the state $|in\rangle$ of Eq. (7), and the post-selected state is

$$|f\rangle = |D_p\rangle \otimes |D_e\rangle, \tag{10}$$

i.e., simultaneous clicks in the detectors D_p and D_e.

They deduce

$$\left[_f(\hat{I}_p \otimes \hat{I}_e)_{in} \right]_w = 0 \tag{11}$$

i.e., vanishing weak value for the simultaneous appearance of the particles in the interacting arms, which is not unreasonable. Moreover

$$\left[_f(\hat{N}_p \otimes \hat{I}_e)_{in} \right]_w = 1 = \left[_f(\hat{N}_e \otimes \hat{I}_p)_{in} \right]_w, \tag{12}$$

which Aharonov *et al.* interpret as implying that there are *two* particle pairs in the apparatus simultaneously: one pair in the N_p- and I_e-arms, the other in the N_e- and I_p-arms. However,

"(q-)uantum mechanics solves the paradox in a remarkable way" by giving

$$\left[_f(\hat{N}_p \otimes \hat{N}_e)_{in} \right]_w = -1, \tag{13}$$

"*i.e.* that there is also *minus* one electron-positron pair in the non-overlapping arms which brings the total down to a single pair." (Quotes are from [20]).

As is evident, the arguments by Aharonov *et al.* rely heavily on an 'ordinary' interpretation of weak values: the weak values of the pair occupation operators are straight-forwardly, and without further ado, identified with the number of pairs in the respective arms of the interferometer.

I question whether this is legitimate. Again, I base this doubt partly on the fact that another choice of the final state $|f\rangle$ than that of Eq. (10) would have given other values for the weak values.

True, the constraints on the choice of the final, post-selected state in this case seems more well-motivated than in the Three-Box Paradox treated above: Given the set-up and the questions asked, it is very natural to choose $|f\rangle = |D_p\rangle \otimes |D_e\rangle$. But nothing fundamental would forbid one to choose other final states also in this case, even if these states might seem somewhat contrived. To illustrate this, I will, however, avoid the slightly more formal complications of the double Mach-Zehnder interferometer of the Hardy set-up and, in the next section, apply my arguments to a single Mach-Zehnder interferometer.

5 A Single, Slightly Generalized Mach-Zehnder Device

Consider then the simple Mach-Zehnder interferometer as illustrated in Fig. 2, but with a more general second beam-splitter BS2. In fact, the most general transformation at a beam-splitter obeying unitarity (= probability conservation) and time-reversal invariance, can be written in obvious matrix notation (and with the same

Fig. 2 Schematic illustration of a single Mach-Zehnder interferometer. (See text to Fig. 1 for explanation of the symbols.) The beam-splitter *BS1* is assumed to be a perfect, 50–50 beam-splitter, while *BS2* is of a more general type (see text, in particular Eq. (14))
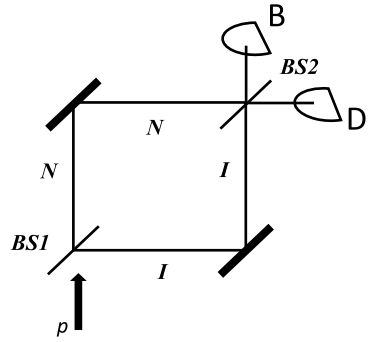
notational conventions for the states as above)

$$\begin{pmatrix} |I\rangle \\ |N\rangle \end{pmatrix} = \begin{pmatrix} q & ir\exp(i\beta) \\ ir\exp(-i\beta) & q \end{pmatrix} \begin{pmatrix} |B\rangle \\ |D\rangle \end{pmatrix}, \tag{14}$$

where q, r and β are real numbers and $q^2 + r^2 = 1$. The successive transitions will then be (the first beam splitter is assumed to be a perfect 50–50 splitter)

$$|p\rangle \xrightarrow{BS1} \frac{1}{\sqrt{2}}\{|N\rangle + i|I\rangle\}$$

$$\xrightarrow{BS2} \frac{i}{\sqrt{2}}\{q + r\exp(-i\beta)\}|B\rangle + \frac{1}{\sqrt{2}}\{q - r\exp(i\beta)\}|D\rangle \tag{15}$$

Now, let the state just before the particle enters the second beam-splitter be the pre-selected state, *i.e.*, choose

$$|in\rangle = \frac{1}{\sqrt{2}}\{|N\rangle + i|I\rangle\}, \tag{16}$$

and one or the other of $|B\rangle$ and $|D\rangle$ as the post-selected state $|f\rangle$. The expressions for the weak value of the respective number operators then read

$$\left[{}_{f=D}(\hat{N})_{in}\right]_w = \frac{q}{q - r\exp(i\beta)} \tag{17}$$

and

$$\left[{}_{f=B}(\hat{N})_{in}\right]_w = \frac{r}{r + q\exp(i\beta)}, \tag{18}$$

with similar results for the weak values of $\hat{I}$.

To say the least, these expressions are difficult to interpret as realistic particle numbers or as ordinary probabilities. Not only would, in such an interpretation, the number of particles in the arms depend on how they are detected, *i.e.*, on the choice of the state $|f\rangle$. Also, the 'number' could take any value, *e.g.*, $[{}_{f=D}(\hat{N})_{in}]_w = -1$ for

$q = 1/\sqrt{5}$, $r = 2/\sqrt{5}$, $\beta = 0$. I think no-one would bet on having minus one particle in the N-arm!

One might argue that one should consider, instead, the real part of these expressions since it is the real part of a weak value that, in the measurement scheme (cf. footnote 1), is most directly related to the pointer position of the meter. But this does not help. Indeed, one finds, *e.g.*,

$$\mathrm{Re}\big[_{f=D}(\hat{N})_{in}\big]_w = \frac{q(q - r\cos\beta)}{1 - 2qr\cos\beta}, \tag{19}$$

which again seem to defy a reasonable interpretation.

These facts cast severe doubts on the whole enterprise of interpreting a weak value in a RSFI fashion as anything reflecting a property of the pre-selected state *per se*.

As a side-remark, it might be interesting to evaluate the corresponding ABL probabilities (see the Appendix below). For example

$$prob\big(\hat{N} = 1 \mid |f\rangle = |D\rangle, |in\rangle\big) = r^2, \tag{20}$$

i.e., the probability of finding the particle in the N-arm, given that it ends up in the D-detector is given by the overlap $|\langle D|N\rangle|^2$ which is not an unreasonable result.

6 Summary

This paper makes two main points:

(1) The weak value is the ratio between two amplitudes. In Sect. 2, I argue that neither experiments nor the basic rules of quantum mechanics can be invoked to motivate a realistic, straight-forward interpretation (an "RSFI") of such a construct.

(2) Moreover, in concrete cases—the Three-Box paradox (Sect. 3), Hardy's paradox (Sect. 4), a slightly generalized single Mach-Zehnder set-up (Sect. 5)—it seems difficult to uphold an RSFI for the weak values of the number operators involved. The main reasons are twofold: the weak value is not a property of the initial state only, but depends also on the choice of the final ('post-selected') state, and non-sensible (complex) values for what is supposed to be a number may easily be obtained.

This casts doubts on any sensible interpretation of weak values representing an 'ordinary' property (see Sect. 2 above for the meaning of 'ordinary' property). In particular, it casts severe doubts on whether the use of weak values really 'explains' the paradoxes—the Three-Box Paradox, Hardy's Paradox—that they were supposed to explain.

Appendix: Relation Between Weak Values and the Aharonov-Bergmann-Lebowitz (ABL) Probability Formula

In 1964—*i.e.*, long before the concept of 'weak measurement + post-selection' was introduced—Aharonov, Bergmann and Lebowitz [23] proved an important result for ordinary, 'strong', measurement. They considered a situation very similar to the one employed in weak measurement: a system is prepared ('pre-selected') in a state $|in\rangle$ and 'post-selected' in a state $|f\rangle$. Then, the ABL conditional probability $prob(o_i \mid |f\rangle, |in\rangle)$ of finding a particular eigenvalue o_i (assumed non-degenerate) of an operator $\hat{O}$ representing an observable O in an ordinary, 'strong' measurement, intermediate in time between the pre- and the post-selection, is

$$\text{ABL:} \quad prob\big(o_i \mid |f\rangle, in\rangle\big) = \frac{|\langle f|o_i\rangle|^2 |\langle o_i|in\rangle|^2}{\sum_j |\langle f|o_j\rangle|^2 |\langle o_j|in\rangle|^2}. \tag{A.1}$$

For the particular case of $\hat{O}$ being a projection/number operator, $\hat{O} = \hat{A} := |a\rangle\langle a|$, onto a particular eigenstate $|a\rangle$ of any (self-adjoint) operator $\hat{\mathbb{A}}$ (the eigenvalues of $\hat{A}$ are 1 and 0, of which the latter could be degenerate even if the former is not; however, that possible degeneracy is also covered by the following treatment), the ABL formula (A.1) reads

$$prob\big(1 \mid |f\rangle, |in\rangle\big) = \frac{|\langle f|\hat{A}|in\rangle|^2}{|\langle f|\hat{A}|in\rangle|^2 + |\langle f|(1-\hat{A})|in\rangle|^2}, \tag{A.2}$$

which is the probability of success in finding $|a\rangle$, and

$$prob\big(0 \mid |f\rangle, |in\rangle\big) = \frac{|\langle f|(1-\hat{A})|in\rangle|^2}{|\langle f|\hat{A}|in\rangle|^2 + |\langle f|(1-\hat{A})|in\rangle|^2}, \tag{A.3}$$

which is the probability of failure in finding it.

In case $\langle f|in\rangle \neq 0$, these expressions may be written in terms of the weak value

$$(_f\hat{A}_{in})_w = \frac{\langle f|\hat{A}|in\rangle}{\langle f|in\rangle} \tag{A.4}$$

as

$$prob\big(1 \mid |f\rangle, |in\rangle\big) = \frac{|(_f\hat{A}_{in})_w|^2}{|(_f\hat{A}_{in})_w|^2 + |[_f(1-\hat{A})_{in}]_w|^2}, \tag{A.5}$$

with a corresponding expression for $prob(0 \mid |f\rangle, |in\rangle)$.[3]

In particular, if $prob(1 \mid |f\rangle, |in\rangle) = 1$ (or $= 0$), *i.e.*, if one is certain to find (respectively not to find) the intermediate eigenstate $|a\rangle$, one also gets the weak value $(_f\hat{A}_{in})_w = 1 (= 0)$ and *vice versa*. Furthermore, if $(_f\hat{A}_{in})_w = -1$, then the ABL probability becomes $prob(1 \mid |f\rangle, |in\rangle) = \frac{1}{5}$.

---

[3]These relations between weak values and ABL probabilities are by no means new; they are given, *e.g.*, in [34, 38].

One may even carry the relation between the ABL probability formula and the weak value one step further: If one assumes that the imaginary part of the weak value vanishes—and this has been the case in all application of weak values to 'paradoxes'—the formula (A.5) may be used to solve for $({}_f\hat{A}_{in})_w$. One finds

$$({}_f\hat{A}_{in})_w = \frac{\sqrt{p}}{\sqrt{p} \pm \sqrt{1-p}}, \quad p := prob\big(1 \mid |f\rangle, |in\rangle\big), \qquad (A.6)$$

where I assume $p \neq \frac{1}{2}$ in case of the minus sign in the $\pm$-ambiguity in the denominator. I note the particular cases

$$prob\big(1 \mid |f\rangle, |in\rangle\big) = 1 \quad \Leftrightarrow \quad ({}_f\hat{A}_{in})_w = 1, \qquad (A.7)$$

as has already been stated, and

$$prob\big(1 \mid |f\rangle, |in\rangle\big) = \frac{1}{5} \quad \Leftrightarrow \quad ({}_f\hat{A}_{in})_w = -1 \quad \text{or} \ = \frac{1}{3}. \qquad (A.8)$$

In sum, there is a very close *numerical* relationship between the ABL probability $prob(1 \mid |f\rangle, |in\rangle)$ of Eq. (A.2) applied to a number operator $\hat{A} = |a\rangle\langle a|$ and the weak value $({}_f\hat{A}_{in})_w$: the ABL probability can be directly expresses in the corresponding weak value, and the weak value, provided it is real, can be expressed in terms of the ABL probability, albeit with a sign ambiguity. In this very restricted sense, there is a correlation between strong and weak pre- and post-selected measurements.

Of course, *conceptually* the two notions are very different.

References

1. Aharonov, Y., Albert, D.Z., Vaidman, L.: How the result of a measurement of a component of the spin of a spin-$\frac{1}{2}$ particle can turn out to be 100. Phys. Rev. Lett. **60**, 1351 (1988)
2. Kofman, A.G., Ashhab, S., Nori, F.: Nonperturbative theory of weak pre- and post-selected measurement. Phys. Rep. **520**, 43 (2012)
3. Svensson, B.E.Y.: Pedagogical review of quantum measurement theory with an emphasis on weak measurements. Quanta **2**(1), 18 (2013). http://quanta.ws
4. Tamir, B., Cohen, E.: Introduction to weak measurements and weak values. Quanta **2**(1), 7 (2013). http://quanta.ws
5. Kocsis, S., Braverman, B., Ravets, S., Stevens, M.J., Mirin, R.P., Shalm, L.K., Steinberg, A.M.: Observing the average trajectories of single photons in a two-slit interferometer. Science **332**, 1170 (2011)
6. Lundeen, J.S., Sutherland, B., Patel, A., Stewart, C., Bambe, C.: Direct measurement of the quantum wave function. Nature **474**, 188 (2011)
7. Ritchie, N.W.M., Story, J.G., Randall, J.G., Hulet, R.G.: Realization of a measurement of a "weak value". Phys. Rev. Lett. **66**, 1107 (1991)
8. Hosten, O., Kwiat, P.: Observation of the spin hall effect of light via weak measurement. Science **319**, 787 (2008)
9. Dixon, P.B., Starling, D.J., Jordan, A.N., Howell, J.C.: Ultrasensitive beam deflection measurement via interferometric weak value amplification. Phys. Rev. Lett. **102**, 173601 (2009)
10. Leggett, A.J., Garg, A.: Quantum mechanics versus macroscopic realism: is the flux there when nobody looks? Phys. Rev. Lett. **54**, 857 (1984)
11. Langford, N.K., Weinhold, T.J., Prevedel, R., Resch, K.J., Gilchrist, A., O'Brien, J.L., Pryde, G.J., White, A.G.: Demonstration of a simple optical gate and its use in bell state analysis. Phys. Rev. Lett. **95**, 210504 (2005)

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

13. Palacios-Laloy, A., Mallet, F., Nguyen, F., Bertet, P., Vion, D., Esteve, D., Korotkov, A.N.: Experimental violation of a Bell's inequality in time with weak measurement. Nat. Phys. **6**, 442 (2010)

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

15. Aharonov, Y., Vaidman, L.: The Two-State Vector Formalism: An Updated Review. Lecture Notes in Physics, vol. 734, Springer, Berlin (2007). arXiv:quant-ph/0105101. See also Aharonov, Y., Popescu, S., Tollaksen, J., A time-symmetric formulation of quantum mechanics, Phys. Today **63** (2010), 27–32 and the ensuing correspondence in Phys. Today **64** (2011), 8–9, 62–63 and October 2011, pp 8–10

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

17. Ravon, T., Vaidman, L.: The three-box paradox revisited. J. Phys. A, Math. Theor. **40**, 2873 (2007)

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

19. Hardy, L.: Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories. Phys. Rev. Lett. **68**, 298 (1992)

20. Aharonov, Y., Botero Popescu, A.S., Reznik, B., Tollaksen, J.: Revisiting Hardy's paradox: counterfactual statements, real measurements, entanglement and weak values. Phys. Lett. A **301**, 130 (2002)

21. Lundeen, J.S., Steinberg, A.M.: Experimental joint weak measurement on a photon pair as a probe of Hardy's paradox. Phys. Rev. Lett. **102**, 020404 (2009)

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

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

24. Vaidman, L.: Weak-measurement element of reality. Found. Phys. **26**, 895 (1996)

25. Vaidman, L.: Counterfactuals in quantum mechanics. In: Greenberger, V., Hentschel, K., Weinert, F. (eds.) Compendium on Quantum Mechanics. Springer, Berlin (2009)

26. Tollaksen, J., Ghoshal, D.: Weak measurements, weak values and entanglement. In: Donkor, E.J., Pirich, A.R., Brandt, H.E. (eds.) Proc. of SPIE, vol. 6573 (2007)

27. Tollaksen, J.: Novel relationship between superoscillations, weak values and modular variables. J. Phys.: Conf. Ser. **70**, 012016 (2007)

28. Dressel, J., Jordan, A.N.: Significance of the imaginary part of the weak value. Phys. Rev. A **85**, 012107 (2012)

29. Leggett, A.J.: Comments on "How the result of a measurement of a components of the spin of a spin-$\frac{1}{2}$ particle can turn out to be 100". Phys. Rev. Lett. **62**, 2325 (1989)

30. Peres, A.: Quantum measurements with postselection. Phys. Rev. Lett. **62**, 2326 (1989)

31. Aharonov, Y., Vaidman, L.: Aharonov and Vaidman reply. Phys. Rev. Lett. **62**, 2327 (1989)

32. Duck, I.M., Stevenson, P.M., Sudarshan, E.C.G.: The sense in which a "weak measurement" of s spin-$\frac{1}{2}$ particle's spin component yields a value 100. Phys. Rev. D, Part. Fields **40**, 2112 (1989)

33. Sokolovski, D.: Are the weak measurements really measurements? Quanta **2**(1), 50 (2013). (http://quanta.ws) and references therein

34. Kastner, R.E.: Weak values and consistent histories in quantum theory. Stud. Hist. Philos. Mod. Phys. **35**, 57 (2004)

35. Steinberg, A.M.: Conditional probabilitites in quantum theory and the tunneling-time controversy. Phys. Rev. A **52**, 32 (1995)

36. Hofmann, H.: What the complex joint probabilities observed in weak measurements can tell us about quantum physics arXiv:1303.0078 (2013) and references therein

37. Rohrlich, D., Aharonov, Y.: Quantum Paradoxes. Wiley, New York (2005)

38. Leavens, C.R., Puerto Gimenez, I., Alonso, D., Sala Mayato, R.: General N-box problem. Phys. Lett. A **359**, 416 (2006)