Pre- and post-selection, weak values and contextuality

2007 J. Phys. A: Math. Theor. 40 9033

(http://iopscience.iop.org/1751-8121/40/30/025)

J. Phys. A: Math. Theor. **40** (2007) 9033–9066 doi:10.1088/1751-8113/40/30/025

Pre- and post-selection, weak values and contextuality

Jeff Tollaksen

Center for Quantum Studies, Department of Physics and Department of Computational and Data Sciences, College of Science, George Mason University, Fairfax, VA 22030, USA

E-mail: jtollaks@gmu.edu

Received 26 July 2006, in final form 23 May 2007
Published 12 July 2007
Online at stacks.iop.org/JPhysA/40/9033

Abstract
By analysing the concept of contextuality (Bell–Kochen–Specker) in terms of pre- and post-selection, it is possible to assign definite values to observables in a new and surprising way. Physical reasons are presented for restrictions on these assignments. When measurements are performed which do not disturb the pre- and post-selection (i.e. weak measurements), then novel *experimental* aspects of contextuality can be demonstrated. We also prove that every PPS-paradox with definite predictions directly implies 'quantum contextuality' which is introduced as the analogue of contextuality at the level of quantum mechanics rather than at the level of hidden variable theories. Finally, we argue that certain results of these measurements (e.g. eccentric weak values outside the eigenvalue spectrum) cannot be explained by a 'classical-like' hidden variable theory.

PACS numbers: 03.65.Ta, 03.65.−w, 06.30.−k

(Some figures in this article are in colour only in the electronic version)

1. Introduction

A traditional concept of the quantum state $|\Psi_{\rm in}\rangle$ is that it generally provides only statistical information about the outcome of an ideal measurement. Therefore, many authors proposed that a theory based on this interpretation of a quantum state could be 'completed' by a hidden variable theory (HVT). Such a theory would bear a similar relationship to quantum mechanics (QM) as classical mechanics has to classical statistical mechanics. That is QM can be understood in terms of a deeper theory, the HVT. The relationship between classical mechanics and classical statistical mechanics is relatively simple because an ideal classical measurement precisely measures a property of a system, *without* affecting the system under study. Measurement of one property will not interfere with the measurement of another property. Thus there is a simple relation between the theory and underlying physical processes (the outcome of measurements directly tells us what values to assign to all variables of the theory).

Moving over to QM, there are two general constraints on any HVT which reproduces QM: (a) the Bell–Kochen–Specker theorem (BKS) [20, 23] and Gleason's [32] theorem showed that any HVT must be contextual, and (b) Bell's theorem [19] showed that any HVT must be non-local. Gleason and BKS proved that one cannot assign unique answers to yes–no questions (posed to single systems) in such a way that measurement simply reveals the answer as a pre-existing property that was intrinsic solely to the quantum system itself. The specification of the HVT is given by a 'value function' $V_{\vec{\psi}}(\hat{A})$ which assigns a value to an observable $\hat{A}$ when an individual system is in the state $\vec{\psi}$. BKS assumed that $V_{\vec{\psi}}(\hat{A})$ should satisfy

$$V_{\vec{\psi}}(F\{\hat{A}\}) = F\{V_{\vec{\psi}}(\hat{A})\}. \tag{1.1}$$

That is, any functional relation of an operator that is a member of a commuting subset of observables must also be satisfied if one substitutes the values for the observables into the functional relations. For example, if a system is characterized by commuting observables $\hat{A}_1$ and $\hat{A}_2$ then condition (1.1) requires that all the *relationships or functions* between these operators should also be satisfied when $V_{\vec{\psi}}(\hat{A}_1)$ and $V_{\vec{\psi}}(\hat{A}_2)$ are substituted into the same functional relations. This condition determines the sum and product rules

$$V_{\vec{\psi}}(\hat{A}_1 + \hat{A}_2) = V_{\vec{\psi}}(\hat{A}_1) + V_{\vec{\psi}}(\hat{A}_2) \tag{1.2}$$

$$V_{\vec{\psi}}(\hat{A}_1\hat{A}_2) = V_{\vec{\psi}}(\hat{A}_1)V_{\vec{\psi}}(\hat{A}_2). \tag{1.3}$$

HVTs which meet these conditions are *noncontextual*: all yes–no questions can be associated with a value assignment $V_{\vec{\psi}}$ which provides a single unique answer, irrespective of the set of other commuting yes–no questions that it is associated with. BKS showed that attempting such an assignment to some observables is inconsistent: in any system (of dimension greater than 2) the 2^n possible 'yes–no' assignments (to the n projection operators representing the yes–no questions) cannot be compatible with the sum (1.2) and product (1.3) rules for all orthogonal resolutions of the identity. For example, consider a complete set of spectral projectors of an operator $\hat{A}$ with discrete eigenvalues, so that $\hat{P}_i = \hat{P}_{A=a_i}$, such that $\sum_{i=1}^{n} V_{\vec{\psi}}(\hat{P}_i) = 1$. Then only one of the projection operators can give a 'yes' assignment ($\hat{P}_i = 1$) and the rest have to be 'no' (i.e. $\hat{P}_i = 0$). However, $\vec{\psi}$ can be decomposed into many different basis sets, and the value that $V_{\vec{\psi}}$ assigns must be independent of the particular basis. BKS showed that this cannot be done.

Given this impossibility proof, an interesting approach to HVTs and BKS is to inquire whether anything new can be learned about experimental situations, in the same spirit as Shimony's apt phrase 'experimental metaphysics'. For example, Bell's theorem led to interesting experiments which tested the notion of whether quantum entanglement could be stronger than classical correlations. Another example is Hardy's paradox [43] (HVTs for position could not be assigned) which was traditionally 'resolved' by arguing that measurements to verify the paradox could not be implemented simultaneously and therefore Hardy's paradox was purely a formal result without empirical consequences. However in [11, 29], we shifted this situation by demonstrating that weak measurements (WM) could be implemented simultaneously on the paradoxical Hardy statements, thereby leading to novel experimental results [35]).

Similarly, the principal result of this paper is to question whether BKS is just a formal result (i.e. negative statements concerning the possibility of a classical-like 'noncontextual HVT') or if BKS has new positive aspects including new experimental consequences. We probe this by utilizing the natural connection between counter-factual statements, pre- and post-selection and the outcome of WMs, the weak value (WV). Although the WVs exhibit strange, non-classical behaviour in BKS situations, they do obey a simple, intuitive and, most importantly, *self-consistent* logic. We then connect contextuality with issues that can be probed experimentally by WMs, thereby distinguishing the instant effort from previous focus on HVTs.

The novel results of this paper lay the ground work for further possible exploitation of the differences between classical and quantum information. By way of example, while the study of the non-classical aspects of entanglement (Einstein–Podolsky–Rosen/Bohm) started as a foundational examination of HVTs, it was subsequently probed experimentally and used as a resource for quantum computation and communication. Similarly, using novel 'generalized' states suggested by TSQM (which are also utilized in the instant paper), Englert *et al* [12] demonstrated a new form of cryptography [18], and experiments based on the optical version of this problem [17] were successfully performed.

1.1. Pre- and post-selection and time-symmetric quantum mechanics

Pre- and post-selection (PPS) were originally probed with the time-symmetric reformulation of quantum mechanics (TSQM, introduced by Aharonov, Bergmann and Lebowitz also known as ABL [2][1], for a good review, see [49, 50]). While TSQM is a new conceptual point of view that has predicted novel, verified effects which *seem* impossible according to standard QM, TSQM is in fact a *reformulation* of QM. Therefore, experiments cannot prove TSQM over QM (or vice versa). The motivation to pursue such reformulations, then, depends on their usefulness. Indeed, we believe that to be useful and interesting, any reformulation of QM should meet several criteria such as those met by TSQM:

- TSQM is consistent with all the predictions made by standard QM;
- TSQM brings out features in QM that were missed before: e.g. the 'weak value' (WV) of an observable which was probed by a new type of quantum measurement called the 'weak measurement' (WM) [5];
- TSQM lead to simplifications in calculations (as occurred with the Feynman reformulation) and stimulated discoveries in other fields: e.g. ABL influenced work in cosmology (e.g. Gell-Mann and Hartle [22]), in superluminal tunnelling ([21, 33]), in quantum information (e.g. the quantum random walk [8] or cryptography [15, 16]), led to the discovery of super-oscillations [51] and new aspects of the uncertainty relations [52], etc, and
- TSQM suggests generalizations of QM that were missed before, e.g. a new solution to the quantum measurement problem.

Using TSQM, we show how to assign definite values to sets of BKS observables in new and surprising ways. We also show how measurement disturbance can arise in a new way when value assignments depend on both the pre- and the post-selection. An 'intriguing' physical reason is presented to explain why this scheme cannot be applied to two or more ideal measurements (IMs): the two IMs interfere with each other because some assignments

[1] ABL is intuitive: $\langle a_j|U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle|^2$ is the probability of obtaining $|a_j\rangle$ having started with $|\Psi_{\rm in}\rangle$. If $|a_j\rangle$ was obtained, then the system collapsed to $|a_j\rangle$ and $|\langle\Psi_{\rm fin}|U_{t\to t_{\rm fin}}|a_j\rangle|^2$ is then the probability of obtaining $|\Psi_{\rm fin}\rangle$. The probability of obtaining $|a_j\rangle$ and $|\Psi_{\rm fin}\rangle$ then is $|\alpha_j|^2$. This is not yet the conditional probability since the post-selection may yield outcomes other than $\langle\Psi_{\rm fin}|$. The probability of obtaining $|\Psi_{\rm fin}\rangle$ is $\sum_j|\alpha_j|^2 = |\langle\Psi_{\rm fin}|\Psi_{\rm in}\rangle|^2 < 1$. The question being investigated concerning probabilities of a_j at t assumes we are successful in obtaining the post-selection and therefore requires the denominator in equation (1.4), $\sum_j|\alpha_j|^2$, which is a renormalization to obtain a proper probability.

of eigenvalues to operators are based on just one of the two boundary conditions (i.e. either the pre- *or* the post-selected vector) while some assignments of eigenvalues are based on both boundary conditions (i.e. both the pre-selected *and* the post-selected vectors, what we call diagonal-PPS). In addition, we show that when measurements are performed which do not disturb the PPS (i.e. WMs), then novel *experimental* aspects of contextuality can be demonstrated. We also prove that every PPS-paradox with definite predictions directly implies 'quantum contextuality' which is introduced as the analogue of contextuality at the level of quantum mechanics rather than at the level of HVTs. For example, we apply the approach developed in this paper to Mermin's work on PPS and contextuality (reviewed in subsubsections 3.2.1–3.2.4). Mermin reflected: '... what follows is not idle theorizing about 'hidden variables'. It is a rock solid quantum mechanical effort to answer a perfectly legitimate quantum mechanical question' [26]. We also demonstrate an isomorphism between WVs in BKS situations and EPR entanglement. Finally, we argue that certain results of WMs (WVs outside the eigenvalue spectrum) cannot be explained by a 'classical-like' HVT.

1.2. The main idea behind time-symmetric quantum mechanics

The main idea behind TSQM focuses on measurements which occur at the present time t while the state is known both at $t_{\rm in} < t$ (past) and at $t_{\rm fin} > t$ (future). More precisely, we start at $t = t_{\rm in}$ with a measurement of a non-degenerate operator $\hat{O}_{\rm in}$ (for simplicity, we consider non-degenerate operators without lack of generality). This yields as one potential outcome the state $|\Psi_{\rm in}\rangle$, i.e. we prepared the 'pre-selected' state $|\Psi_{\rm in}\rangle$. At the later time $t_{\rm fin}$, we perform another measurement of a non-degenerate operator $\hat{O}_{\rm fin}$ which yields one possible outcome: the post-selected state $|\Psi_{\rm fin}\rangle$. At an intermediate time $t \in [t_{\rm in}, t_{\rm fin}]$, we measure a non-degenerate observable $\hat{A}$, with eigenvectors $\{|a_j\rangle\}$. We wish to determine the conditional probability of a_j, given that we have both boundary conditions, $|\Psi_{\rm in}\rangle$ and $\langle\Psi_{\rm fin}|$. To answer this, we use the time displacement operator: $U_{t_{\rm in}\to t} = \exp\{-{\rm i}H(t - t_{\rm in})\}$, where H is the Hamiltonian for the free system (for simplicity, we assume H is time independent and set $\hbar = 1$). The standard theory of collapse states that the system collapses into an eigenstate $|a_j\rangle$ *after* the measurement at t with an amplitude $\langle a_j|U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle$. The amplitude for our series of events including the post-selection is $\alpha_j \equiv \langle\Psi_{\rm fin}|U_{t\to t_{\rm fin}}|a_j\rangle\langle a_j|U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle$ illustrated in figure 1(a). This means that the conditional probability of measuring a_j given $|\Psi_{\rm in}\rangle$ is pre-selected and $|\Psi_{\rm fin}\rangle$ will be post-selected is

$$\mathrm{Prob}(a_j, t|\Psi_{\rm in}, t_{\rm in}; \Psi_{\rm fin}, t_{\rm fin}) = \frac{\left|\langle\Psi_{\rm fin}|U_{t\to t_{\rm fin}}|a_j\rangle\langle a_j|U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle\right|^2}{\sum_n \left|\langle\Psi_{\rm fin}|U_{t\to t_{\rm fin}}|a_n\rangle\langle a_n|U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle\right|^2} \qquad (1.4)$$

which is the ABL formula [2]. As a first step towards understanding the underlying time-symmetry in the ABL formula, we consider the time reverse of the numerator of equation (1.4). First we apply $U_{t\to t_{\rm fin}}$ on $\langle\Psi_{\rm fin}|$ instead of on $\langle a_j|$. We note that $\langle\Psi_{\rm fin}|U_{t\to t_{\rm fin}} = \left\langle U^\dagger_{t\to t_{\rm fin}}\Psi_{\rm fin}\right|$ by using the well-known QM symmetry $U^\dagger_{t\to t_{\rm fin}} = \{{\rm e}^{-{\rm i}H(t_{\rm fin}-t)}\}^\dagger = {\rm e}^{{\rm i}H(t_{\rm fin}-t)} = {\rm e}^{-{\rm i}H(t-t_{\rm fin})} = U_{t_{\rm fin}\to t}$. We also apply $U_{t_{\rm in}\to t}$ on $\langle a_j|$ instead of on $|\Psi_{\rm in}\rangle$ which yields the time-reverse reformulation of the numerator of equation (1.4), $\left\langle U_{t_{\rm fin}\to t}\Psi_{\rm fin}\middle|a_j\right\rangle\!\left\langle U_{t\to t_{\rm in}}a_j\middle|\Psi_{\rm in}\right\rangle$ as depicted in figure 1(b). Further work is needed to formulate what we mean by time-symmetry with respect to boundary conditions, otherwise referred to as the 2-vectors in TSQM. For example, if we are interested in the probability for possible outcomes of a_j at time t, we must consider both $U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle$ and $\left\langle U_{t_{\rm fin}\to t}\Psi_{\rm fin}\right|$, since these expressions propagate the pre- and post-selection to the present time t (see the conjunction of both figures 1(a) and (b) giving 1(c); these 2-vectors

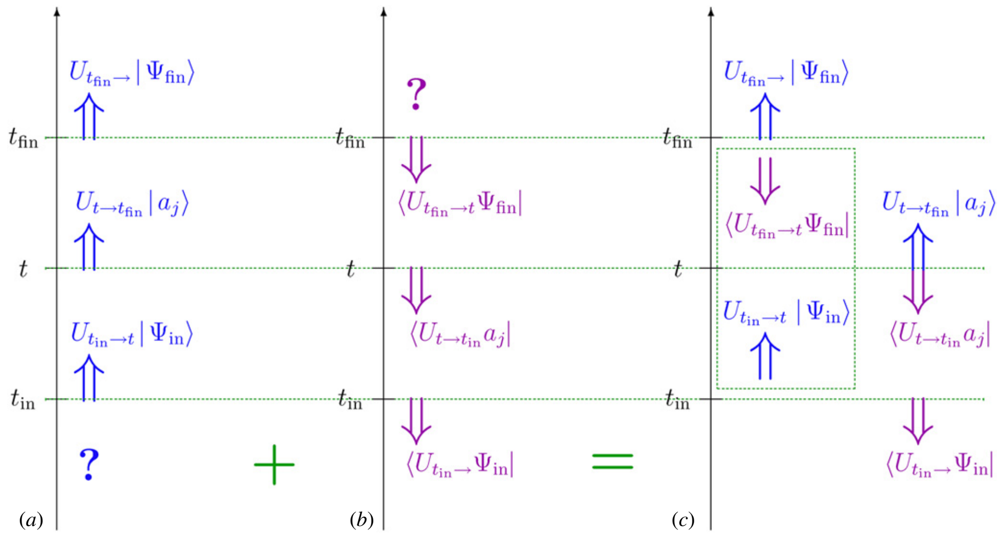

Figure 1. Time-reversal symmetry in probability amplitudes.

are not just the time reverse of each other). This represents the basic idea behind TSQM as can be seen in the re-expression of equation (1.4):

$$\mathrm{Prob}(a_j, t|\Psi_{\rm in}, t_{\rm in}; \Psi_{\rm fin}, t_{\rm fin}) = \frac{\left|\langle U_{t_{\rm fin}\to t}\Psi_{\rm fin}|a_j\rangle\langle a_j|U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle\right|^2}{\sum_n \left|\langle U_{t_{\rm fin}\to t}\Psi_{\rm fin}|a_n\rangle\langle a_n|U_{t_{\rm in}\to t}|\Psi_{\rm in}\rangle\right|^2}. \qquad (1.5)$$

While this mathematical manipulation clearly proves that TSQM is consistent with QM, it yields a very different interpretation. For example, the action of $U_{t_{\rm fin}\to t}$ on $\langle\Psi_{\rm fin}|$ (i.e. $\left\langle U_{t_{\rm fin}\to t}\Psi_{\rm fin}\right|$) can be interpreted to mean that the time displacement operator $U_{t_{\rm fin}\to t}$ sends $\langle\Psi_{\rm fin}|$ back in time from the time $t_{\rm fin}$ to the present t.

1.3. Pre- and post-selection and the three-box paradox

To illustrate some of the surprising consequences of TSQM (which will be relevant to contextuality), we consider the three-box paradox ([3], verified experimentally [34]) which uses a *single* quantum particle that is placed in a superposition of being in three closed, separated boxes. The particle is pre-selected to be in the state $|\Psi_{\rm in}\rangle = 1/\sqrt{3}(|A\rangle + |B\rangle + |C\rangle)$, where $|A\rangle$, $|B\rangle$ and $|C\rangle$ denote the particle localized in boxes A, B, or C, respectively. The particle is post-selected to be in the state $|\Psi_{\rm fin}\rangle = 1/\sqrt{3}(|A\rangle+|B\rangle-|C\rangle)$. If an IM is performed on box A in the intermediate time (e.g. we open the box), then the particle is found in box A with certainty. This is confirmed by the ABL [2] probability for projection in A:

$$\mathrm{Prob}(\hat{\mathbf{P}}_A) = \frac{|\langle\Psi_{\rm fin}|\hat{\mathbf{P}}_A|\Psi_{\rm in}\rangle|^2}{|\langle\Psi_{\rm fin}|\hat{\mathbf{P}}_A|\Psi_{\rm in}\rangle|^2 + |\langle\Psi_{\rm fin}|\hat{\mathbf{P}}_B + \hat{\mathbf{P}}_C|\Psi_{\rm in}\rangle|^2} = 1.$$

This can also be seen intuitively by contradiction: suppose we do not find the particle in box$|A\rangle$. In that case, since we do not interact with box $|B\rangle$ or $|C\rangle$, we would have to conclude that the state that remains after we did not find it in $|A\rangle$ is proportional to $|B\rangle + |C\rangle$. But this is orthogonal to the post-selection (which we know will definitely be obtained). Because this is a contradiction, we conclude that the particle must be found in box A. Similarly, using the same reasoning, the particle can be found with certainty in box B, i.e. $\mathrm{Prob}(\hat{\mathbf{P}}_B = 1) = 1$.

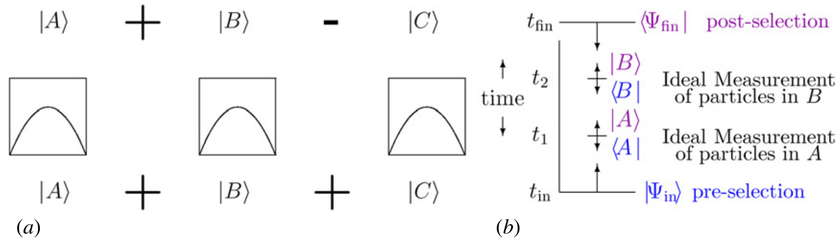

Figure 2. (*a*) Pre-selected vector $|\Psi_{\mathrm{in}}\rangle = 1/\sqrt{3}(|A\rangle + |B\rangle + |C\rangle)$ propagates forwards in time from t_{in} to t_1, and post-selected vector $|\Psi_{\mathrm{fin}}\rangle = 1/\sqrt{3}(|A\rangle + |B\rangle - |C\rangle)$ propagating backwards in time from t_{fin} to t_2. (*b*) Ideal measurement of $\hat{\mathbf{P}}_A$ at t_1 and of $\hat{\mathbf{P}}_B$ at t_2.

The 'paradox' can thus be stated: 'given that we only have a single particle, in what sense, if any, can these two definite statements (i.e. $\mathrm{Prob}(\hat{\mathbf{P}}_A = 1) = 1$ and $\mathrm{Prob}(\hat{\mathbf{P}}_B = 1) = 1$) be *simultaneously* true'?

1.4. Counterfactuals

There is a widespread tendency to 'resolve' these paradoxes by pointing out that there is an element of counter-factual reasoning in them: the contradictions arise only because inferences are made that do not refer to actual experiments. Had the experiment actually been performed, then standard measurement theory predicts that the system would have been disrupted so that no paradoxical implications arise. By way of example, suppose we applied this to the three-box paradox: the resolution then is that there is no meaning to say that the particle is in both boxes without actually *measuring* both boxes during the intermediate time. That is $\mathrm{Prob}(\hat{\mathbf{P}}_A = 1) = 1$ if only box A is opened, while $\mathrm{Prob}(\hat{\mathbf{P}}_B = 1) = 1$ if only box B is opened. If IMs are performed on *both* boxes A *and* B, then obviously the particle will not be found in both boxes, i.e. $\hat{\mathbf{P}}_A\hat{\mathbf{P}}_B = 0$, and the paradox disappears!

While we cannot discern the sense in which $\mathrm{Prob}(\hat{\mathbf{P}}_A = 1) = 1$ is *simultaneously* true with $\mathrm{Prob}(\hat{\mathbf{P}}_B = 1) = 1$ by using IMs, there is still a mystery here: even though $\hat{\mathbf{P}}_A$ and $\hat{\mathbf{P}}_B$ commute with each other, measurement of one can disturb the measurement of the other. A novel explanation of this situation, we suggest, is as follows.

- In order to deduce $\mathrm{Prob}(\hat{\mathbf{P}}_A = 1) = 1$ (or $\mathrm{Prob}(\hat{\mathbf{P}}_B = 1) = 1$), ABL requires information from both the pre- *and* the post-selected vectors. We call this situation 'diagonal-pre- and post-selections' or 'diagonal-PPS'.
- When we actually measure $\hat{\mathbf{P}}_A$, then this measurement will limit the 'propagation' of *both* the pre- and the post-selected vectors that are required for any attempt to ascertain other intermediate values, such as $\hat{\mathbf{P}}_B$ (see figure 2(*b*)).
- If we subsequently were to measure $\hat{\mathbf{P}}_B$, then the necessary information from both the pre- and post-selected vectors is no longer available (i.e. information from t_{in} cannot propagate beyond the IM of $\hat{\mathbf{P}}_A$ at time t_1 due to the disturbance caused by the IM of $\hat{\mathbf{P}}_A$).
- *Thus, even though* $\hat{\mathbf{P}}_A$ *and* $\hat{\mathbf{P}}_B$ *commute, the IM of one can disturb the IM of the other: a violation of the product rule.*[2]

[2] In general, if $|\Psi_1\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue a and $|\Psi_2\rangle$ is an eigenvector of $\hat{B}$ with eigenvalue b and $[\hat{A}, \hat{B}] = 0$, then if $\hat{A}$ and $\hat{B}$ are known only by either pre-selection *or* post-selection, then the product rule is valid, i.e. $\hat{A}\hat{B} = ab$. However if $\hat{A}$ and $\hat{B}$ are known by both pre-selection *and* post-selection, then the product rule is not generally valid, i.e. $\hat{A}\hat{B} \neq ab$, i.e. they can still disturb each other, even though they commute [28].

In addition to this insight, we have proven [11, 29] that one should not be so quick in throwing away counter-factual reasoning; though indeed counter-factual statements have no observational meaning, such reasoning is actually a very good pointer towards interesting physical situations. *Without invoking counter-factual reasoning*, we have shown that the apparently paradoxical reality implied counter-factually has new, *experimentally accessible* consequences. These observable consequences become evident in terms of WVs and WMs, which allow us to test—to some extent—assertions that have been otherwise regarded as counter-factual.

Again, the main argument against counter-factual statements is that if we actually perform IMs to test them, we disturb the system significantly, and such disturbed conditions hide the counter-factual situation, so no paradox arises. With the disturbance-based understanding as to why both statements are not simultaneously true, we can now see the 'sense' in which the definite ABL assignments should be simultaneously relevant. Our main argument is that if one does not perform absolutely precise (ideal) measurements but one with finite accuracy, then one can bound the disturbance on the system. For example, according to Heisenberg's uncertainty relations, a precise measurement of position reduces the uncertainty in position to zero $\Delta x = 0$ but produces an infinite uncertainty in momentum $\Delta p = \infty$. On the other hand, if we measure the position only up to some finite precision $\Delta x = \Delta$ we can limit the disturbance of momentum to a finite amount $\Delta p \geqslant \hbar/\Delta$. By replacing precise measurements with a bounded-measurement paradigm, e.g. WMs, counter-factual thought experiments become experimentally accessible.

1.5. Weak values and weak measurements

ABL considered measurements *between* two successive IMs where the transition from a pre-selected state $|\Psi_{\rm in}\rangle$ to a post-selected state $|\Psi_{\rm fin}\rangle$ is generally disturbed by the IM performed during the intermediate time. A subsequent theoretical development arising out of the ABL work was the introduction of the 'weak value' (WV) of an observable which can be probed by a new type of measurement called the 'weak measurement' (WM) [5]. Part of the motivation behind these measurements is to explore the relationship between $|\Psi_{\rm in}\rangle$ and $|\Psi_{\rm fin}\rangle$ by reducing the disturbance on the system at the intermediate time. This is useful in many ways, e.g. if a WM of $\hat{A}$ is performed at the intermediate time $t \in [t_{\rm in}, t_{\rm fin}]$ then, in contrast to the ABL situation, the basic object in the entire interval $t_{\rm in} \to t_{\rm fin}$ for the purpose of calculating *other* WVs for other measurements is the pair of states $|\Psi_{\rm in}\rangle$ and $|\Psi_{\rm fin}\rangle$.

1.5.1. Quantum measurements. WMs [5] originally grew out of the quantum measurement theory developed by von Neumann [37][3]. First we consider IMs of observable $\hat{A}$ by using an interaction Hamiltonian $H_{\rm int}$ of the form $H_{\rm int} = -\lambda(t)\hat{Q}_{\rm md}\hat{A}$, where $\hat{Q}_{\rm md}$ is an observable of the measuring device (e.g. the position of the pointer) and $\lambda(t)$ is a coupling constant which determines the duration and strength of the measurement. For an impulsive measurement we need the coupling to be strong and short and thus take $\lambda(t) \neq 0$ only for $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ and set $\lambda = \int_{t_0-\varepsilon}^{t_0+\varepsilon} \lambda(t)\,{\rm d}t$. We may then neglect the time evolution given by $H_{\rm s}$ and $H_{\rm md}$ in the complete Hamiltonian $H = H_{\rm s} +_{\rm md} +H_{\rm int}$. Using the Heisenberg equations of motion for the momentum $\hat{P}_{\rm md}$ of the measuring device (conjugate to the position $\hat{Q}_{\rm md}$), we see that $\hat{P}_{\rm md}$ evolves according to $\frac{{\rm d}\hat{P}_{\rm md}}{{\rm d}t} = \lambda(t)\hat{A}$. Integrating this, we see that $P_{\rm md}(T) - P_{\rm md}(0) = \lambda\hat{A}$, where $P_{\rm md}(0)$ characterizes the initial state of the measuring device and $P_{\rm md}(T)$ characterizes

[3] WMs and their outcome, WVs, can be derived in all approaches to quantum measurement theory. For example, the usual projective measurement typically utilized in quantum experiments is a special case of these WMs [45].

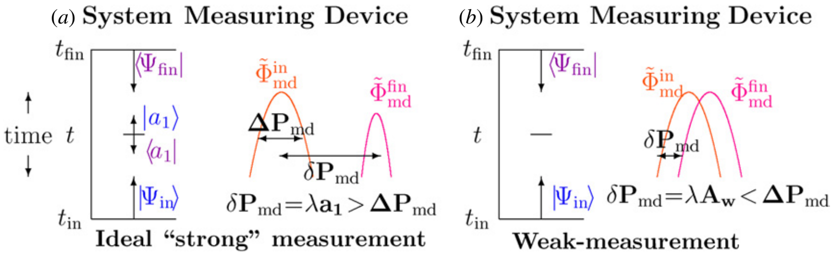

Figure 3. (*a*) With an ideal or 'strong' measurement at *t* (characterized e.g. by $\delta P_{\rm md} = \lambda a_1 \gg \Delta P_{\rm md}$), then ABL gives the probability of obtaining a collapse onto eigenstate a_1 by propagating $\langle\Psi_{\rm fin}|$ backwards in time from $t_{\rm fin}$ to t and $|\Psi_{\rm in}\rangle$ forwards in time from $t_{\rm in}$ to t; in addition, the collapse caused by IM at t creates a new boundary condition $|a_1\rangle\,\langle a_1|$ at time $t \in [t_{\rm in}, t_{\rm fin}]$. (*b*) If a WM is performed at t (characterized e.g. by $\delta P_{\rm md} = \lambda A_w \ll \Delta P_{\rm md}$), then the outcome of the WM, the WV, can be calculated by propagating the state $\langle\Psi_{\rm fin}|$ backwards in time from $t_{\rm fin}$ to t and the state $|\Psi_{\rm in}\rangle$ forwards in time from $t_{\rm in}$ to t; the WM does not cause a collapse and thus no new boundary condition is created at time t.

the final state. To make a more precise determination of $\hat{A}$ requires that the shift in $P_{\rm md}$, i.e. $\delta P_{\rm md} = P_{\rm md}(T) - P_{\rm md}(0)$, be distinguishable from its uncertainty, $\Delta P_{\rm md}$. This occurs, e.g., if $P_{\rm md}(0)$ and $P_{\rm md}(T)$ are more precisely defined and/or if λ is sufficiently large (see figure 3(*a*)). However, under these conditions (e.g. if the measuring device approaches a delta function in $P_{\rm md}$), then the disturbance or back-reaction on the system is increased due to a larger $H_{\rm int}$, the result of the larger $\Delta Q_{\rm md}$ $\left(\Delta Q_{\rm md} \geqslant \frac{1}{\Delta P_{\rm md}}\right)$. When $\hat{A}$ is measured in this way, then any operator $\hat{O}$ ($[\hat{A}, \hat{O}] \neq 0$) is disturbed because it evolved according to $\frac{\rm d}{{\rm d}t}\hat{O} = {\rm i}\lambda(t)[\hat{A}, \hat{O}]\hat{Q}_{\rm md}$, and since $\lambda\Delta Q_{\rm md}$ is not zero, $\hat{O}$ changes in an uncertain way proportional to $\lambda\Delta Q_{\rm md}$.

In the Schroedinger picture, the time evolution operator for the complete system from $t = t_0 - \varepsilon$ to $t = t_0 + \varepsilon$ is $\exp\left\{-{\rm i}\int_{t_0-\varepsilon}^{t_0+\varepsilon} H(t)\,{\rm d}t\right\} = \exp\{-{\rm i}\lambda\hat{Q}_{\rm md}\hat{A}\}$. This shifts $P_{\rm md}$ (see figure 3(*a*)). If before the measurement the system was in a superposition of eigenstates of $\hat{A}$, then the measuring device will also be superposed proportional to the system. This leads to the 'quantum measurement problem'. A conventional solution to this problem is to argue that because the measuring device is macroscopic, it cannot be in a superposition, and so it will 'collapse' into one of these states and the system will collapse with it.

1.5.2. Weakening the interaction between system and measuring device. Following our intuition (subsection 1.4) concerning disturbance and counter-factual statements, we now perform measurements which do not disturb the system. The interaction $H_{\rm int} = -\lambda(t)\hat{Q}_{\rm md}\hat{A}$ is weakened by minimizing $\lambda\Delta Q_{\rm md}$. For simplicity, we consider $\lambda \ll 1$ (assuming without lack of generality that the state of the measuring device is a Gaussian with spreads $\Delta P_{\rm md} = \Delta Q_{\rm md} = 1$). We set ${\rm e}^{-{\rm i}\lambda\hat{Q}_{\rm md}\hat{A}} \approx 1 - {\rm i}\lambda\hat{Q}_{\rm md}\hat{A}$ and consider a theorem.

Theorem 1.

$$\hat{A}|\Psi\rangle = \langle\hat{A}\rangle|\Psi\rangle + \Delta A|\Psi_\perp\rangle, \tag{1.6}$$

where $\langle\hat{A}\rangle = \langle\Psi|\hat{A}|\Psi\rangle$, $|\Psi\rangle$ *is any vector in Hilbert space,* $\Delta A^2 = \langle\Psi|(\hat{A} - \langle\hat{A}\rangle)^2|\Psi\rangle$, *and* $|\Psi_\perp\rangle$ *is a state such that* $\langle\Psi|\Psi_\perp\rangle = 0$.

Proof. To prove this theorem, we begin with $A|\psi\rangle = \langle A\rangle|\psi\rangle + A|\psi\rangle - \langle A\rangle|\psi\rangle$ now, we set $|\widetilde{\psi}_\perp\rangle = A|\psi\rangle - \langle A\rangle|\psi\rangle$, so $\langle\widetilde{\psi}_\perp|\psi\rangle = (\langle\psi|A - \langle\psi|\langle A\rangle)|\psi\rangle = \langle\psi|A|\psi\rangle - \langle A\rangle\langle\psi|\psi\rangle = 0$ now we set $|\psi_\perp\rangle = b|\widetilde{\psi}_\perp\rangle$, where $|\psi_\perp\rangle$ is normalized and b real (note that $\langle\psi|\psi_\perp\rangle = 0$). so $A|\psi\rangle = \langle A\rangle|\psi\rangle + b|\psi_\perp\rangle$. Now we multiply from the left by $\langle\psi_\perp|$, and we get $\langle\psi_\perp|A|\psi\rangle = b$.

Now we can see that $\langle\psi|A^2|\psi\rangle = \langle\psi|A(\langle A\rangle|\psi\rangle + b|\psi_\perp\rangle) = \langle\psi|(\langle A\rangle^2|\psi\rangle + b\langle A\rangle|\psi_\perp\rangle + bA|\psi_\perp\rangle) = \langle A\rangle^2 + b\langle\psi|A|\psi_\perp\rangle$ so $\langle A^2\rangle - \langle A\rangle^2 = b\langle\psi|A|\psi_\perp\rangle = b^2$ which means that $b = \sqrt{\langle A^2\rangle - \langle A\rangle^2} = \Delta A$ and the result $A|\psi\rangle = \langle A\rangle|\psi\rangle + \Delta A|\psi_\perp\rangle$ is proved. $\square$

Using this, we can then show that before the post-selection, the system state is

$$\mathrm{e}^{-\mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\hat{A}}|\Psi_{\mathrm{in}}\rangle = (1 - \mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\hat{A})|\Psi_{\mathrm{in}}\rangle = (1 - \mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\langle\hat{A}\rangle)|\Psi_{\mathrm{in}}\rangle - \mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\Delta\hat{A}|\Psi_{\mathrm{in}\perp}\rangle \tag{1.7}$$

Using the norm of this state $\|(1 - \mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\hat{A})|\Psi_{\mathrm{in}}\rangle\|^2 = 1 + \lambda^2\hat{Q}_{\mathrm{md}}^2\langle\hat{A}^2\rangle$, the probability of leaving $|\Psi_{\mathrm{in}}\rangle$ unchanged after the measurement is

$$\frac{1 + \lambda^2\hat{Q}_{\mathrm{md}}^2\langle\hat{A}\rangle^2}{1 + \lambda^2\hat{Q}_{\mathrm{md}}^2\langle\hat{A}^2\rangle} \longrightarrow 1 \qquad (\lambda \to 0), \tag{1.8}$$

while the probability of disturbing the state (i.e. of obtaining $|\Psi_{in\perp}\rangle$) is

$$\frac{\lambda^2\hat{Q}_{\mathrm{md}}^2\Delta\hat{A}^2}{1 + \lambda^2\hat{Q}_{\mathrm{md}}^2\langle\hat{A}^2\rangle} \longrightarrow 0 \qquad (\lambda \to 0) \tag{1.9}$$

The final state of the measuring device is now a superposition of many substantially overlapping Gaussians with probability distribution given by $\mathrm{Prob}(P_{\mathrm{md}}) = \sum_i |\langle a_i|\Psi_{\mathrm{in}}\rangle|^2 \exp\left\{-\frac{(P_{\mathrm{md}}-\lambda a_i)^2}{2\Delta P_{\mathrm{md}}^2}\right\}$. This sum is a Gaussian mixture, so it can be approximated by a single Gaussian $\tilde{\Phi}_{\mathrm{md}}^{\mathrm{fin}}(P_{\mathrm{md}}) \approx \langle P_{\mathrm{md}}|\,\mathrm{e}^{-\mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\langle\hat{A}\rangle}\big|\Phi_{\mathrm{md}}^{\mathrm{in}}\big\rangle \approx \exp\left\{-\frac{(P_{\mathrm{md}}-\lambda\langle\hat{A}\rangle)^2}{\Delta P_{\mathrm{md}}^2}\right\}$ centred on $\lambda\langle\hat{A}\rangle$.

1.5.3. Adding a post-selection to the weakened interaction: weak values and weak measurements. Appendix A and [54] show how to obtain information from a quantum measurement without causing a disturbance to the state. Having thereby established a new measurement paradigm—information gain without disturbance—it is fruitful to inquire whether this type of measurement reveals new values or properties. With WMs (which involve adding a post-selection to this ordinary—but weakened—von Neumann measurement), the measuring device registers a new value, the WV. As an indication of this, we insert a complete set of states $\{|\Psi_{\mathrm{fin}}\rangle_j\}$ into the outcome of the weak interaction of subsubsection 1.5.2, i.e. into the expectation value $\langle\hat{A}\rangle$:

$$\langle\hat{A}\rangle = \langle\Psi_{\mathrm{in}}|\left[\sum_j |\Psi_{\mathrm{fin}}\rangle_j\langle\Psi_{\mathrm{fin}}|_j\right]\hat{A}|\Psi_{\mathrm{in}}\rangle = \sum_j |\langle\Psi_{\mathrm{fin}}|_j\Psi_{\mathrm{in}}\rangle|^2\frac{\langle\Psi_{\mathrm{fin}}|_j\hat{A}|\Psi_{\mathrm{in}}\rangle}{\langle\Psi_{\mathrm{fin}}|_j\Psi_{\mathrm{in}}\rangle}. \tag{1.10}$$

If we interpret the states $|\Psi_{\mathrm{fin}}\rangle_j$ as the outcomes of a final IM on the system (i.e. a post-selection) then performing a WM (e.g. with $\lambda\Delta Q_{\mathrm{md}} \to 0$) during the intermediate time $t \in [t_{\mathrm{in}}, t_{\mathrm{fin}}]$ provides the coefficients for $|\langle\Psi_{\mathrm{fin}}|_j\Psi_{\mathrm{in}}\rangle|^2$ which gives the probabilities $\mathrm{Prob}(j)$ for obtaining a pre-selection of $\langle\Psi_{\mathrm{in}}|$ and a post-selection of $|\Psi_{\mathrm{fin}}\rangle_j$. The intermediate WM does not disturb these states and the quantity $A_{\mathrm{w}}(j) \equiv \frac{\langle\Psi_{\mathrm{fin}}|_j\hat{A}|\Psi_{\mathrm{in}}\rangle}{\langle\Psi_{\mathrm{fin}}|_j\Psi_{\mathrm{in}}\rangle}$ is the WV of $\hat{A}$ given a particular final post-selection $\langle\Psi_{\mathrm{fin}}|_j$. Thus, from the definition $\langle\hat{A}\rangle = \sum_j \mathrm{Prob}(j)A_{\mathrm{w}}(j)$, one can think of $\langle\hat{A}\rangle$ for the whole ensemble as being constructed out of sub-ensembles of pre- and post-selected states in which the WV is multiplied by a probability for a post-selected state.

Indeed, the WV arises naturally from a weakened measurement with post-selection: taking $\lambda \ll 1$, the final state of measuring device in the momentum representation becomes

$$\begin{aligned}
\langle P_{\mathrm{md}}|\langle\Psi_{\mathrm{fin}}|\,\mathrm{e}^{-\mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\hat{A}}|\Psi_{\mathrm{in}}\rangle\big|\Phi_{\mathrm{in}}^{\mathrm{MD}}\big\rangle &\approx \langle P_{\mathrm{md}}|\langle\Psi_{\mathrm{fin}}|1+\mathrm{i}\lambda\hat{Q}_{\mathrm{md}}\hat{A}|\Psi_{\mathrm{in}}\rangle\big|\Phi_{\mathrm{in}}^{\mathrm{MD}}\big\rangle \\
&\approx \langle P_{\mathrm{md}}|\langle\Psi_{\mathrm{fin}}|\Psi_{\mathrm{in}}\rangle\left\{1+\mathrm{i}\lambda\hat{Q}\frac{\langle\Psi_{\mathrm{fin}}|\hat{A}|\Psi_{\mathrm{in}}\rangle}{\langle\Psi_{\mathrm{fin}}|\Psi_{\mathrm{in}}\rangle}\right\}\big|\Phi_{\mathrm{in}}^{\mathrm{MD}}\big\rangle \\
&\approx \langle\Psi_{\mathrm{fin}}|\Psi_{\mathrm{in}}\rangle\langle P_{\mathrm{md}}|\,\mathrm{e}^{-\mathrm{i}\lambda\hat{Q}A_{\mathrm{w}}}\big|\Phi_{\mathrm{in}}^{\mathrm{MD}}\big\rangle \\
&\to \langle\Psi_{\mathrm{fin}}|\Psi_{\mathrm{in}}\rangle\exp\{-(P_{\mathrm{md}}-\lambda A_{\mathrm{w}})^2\}, \\
&\quad \text{where}\quad A_{\mathrm{w}}=\frac{\langle\Psi_{\mathrm{fin}}|\hat{A}|\Psi_{\mathrm{in}}\rangle}{\langle\Psi_{\mathrm{fin}}|\Psi_{\mathrm{in}}\rangle}.
\end{aligned} \tag{1.11}$$

The final state of the measuring device is almost un-entangled with the system and is shifted by a very unusual quantity, the WV, A_{w}, which is not in general an eigenvalue of $\hat{A}$. We have used such limited disturbance measurements to explore many paradoxes (see, e.g. [11, 29]). A number of experiments have been performed to test the predictions made by WMs and results have proven to be in very good agreement with theoretical predictions [55–59]. Since eigenvalues or expectation values can be *derived* from WVs [1], we believe that the WV is indeed of fundamental importance in QM. In addition, the WV is the relevant quantity for all generalized weak interactions with an environment, not just measurement interactions. The only requirement being that the 2-vectors, i.e. the pre- and post-selection, are not significantly disturbed by the environment.

2. Pre- and post-selected paradoxes and contextuality

Now that we have completed our review, we will spend much of this paper showing how to utilize pre- and post-selection to assign definite values to observables in new and surprising ways. We do this for systems that, from the perspective of HVTs, are 'primed to exhibit' contextuality. We first show that these new assignments can be verified individually by performing IMs. While this assignment suggests novel connections between what could be said about the state before the IM and after, in general, the IM creates a disturbance and thus creates an uncertain relationship between the state before and after, reflecting Bub and Brown's [44] concern 'ensembles which have been pre-selected and post-selected via an arbitrary intervening measurement . . . are not well defined without specification of the intervening measurement'. Indeed, we show [29] how measurement disturbance can arise in new ways when IMs are performed on PPS systems, even, surprisingly, for commuting observables.

However, in contrast to IMs, ensembles that are pre- and post-selected are well defined for any WM or WV, and therefore become useful for probing contextuality at an empirical level. We use such limited disturbance measurements (WMs) to experimentally test the paradoxes implied by counter-factual statements in situations involving contextuality. For example, we apply this new approach to several specific examples (the three-box paradox subsection 2.1, Mermin's 4D example section 3, GHZ section 4) as well as a general proof for all logical pre- and post-selected paradoxes subsection 2.2. Although the outcomes of WMs suggest a story which appears to be even stranger than the original one (non-classical WVs, etc), the situation is in fact far better. WVs obey a simple, intuitive and, most importantly, *self-consistent* logic. This contrasts with the logic of the original counter-factual statements which is not internally self-consistent and leads into paradoxes.

2.1. Weak values and the three-box paradox

Before considering the general proofs, we return first to the three-box paradox. In subsection 1.3, we discussed eight different vectors that were relevant for constructing the paradox: the pre-selection $|\Psi_{\text{in}}\rangle = 1/\sqrt{3}(|A\rangle + |B\rangle + |C\rangle)$, the post-selection $|\Psi_{\text{fin}}\rangle = 1/\sqrt{3}(|A\rangle + |B\rangle - |C\rangle)$, and six possible intermediate measurements which give rise to the paradoxical inferences. The BKS conditions discussed in the introduction, namely the fact that $\vec{\psi}$ can be decomposed into many different basis sets, and the value that $V_{\vec{\psi}}$ assigns must be independent of the particular basis, can be represented pictorially by assigning 'no's (black circles), 'yes's (white circles) and connecting lines as orthogonality relations. Leifer and Spekkens demonstrated a novel direct connection between each of the eight vectors in the three-box paradox and the eight vectors in the Clifford/BKS-proof of contextuality: it is readily seen ([25]'s figure) that no HVT assignment can be made that is consistent with the orthogonality relations, since in order for noncontextuality to hold, no two orthogonal pairs can be white, which is violated by the central two circles.

Because IMs are used, the six possible measurements during the intermediate time are 'treated as counter-factual alternatives in the proof of contextuality' [24]. Although these have no physical meaning, they are a good guide to interesting physical situations in terms of WVs and WMs. For example, we can calculate the WVs of the number of particles in each box:

$$(|A\rangle\langle A|)_w = \frac{\langle\Psi_{\text{fin}}|A\rangle\langle A|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{fin}}\rangle} = \frac{\frac{1}{\sqrt{3}}\{\langle A| + \langle B| - \langle C|\}|A\rangle\langle A|\frac{1}{\sqrt{3}}\{|A\rangle + |B\rangle + |C\rangle\}}{\frac{1}{\sqrt{3}}\{\langle A| + \langle B| - \langle C|\}\frac{1}{\sqrt{3}}\{|A\rangle + |B\rangle + |C\rangle\}}$$
$$= \frac{\frac{1}{3}1\cdot 1}{\frac{1}{3}(1+1-1)} = 1.$$

We can more easily ascertain the WVs without calculation due to the following theorems.

Theorem 2. *The sum of the WVs is equal to the WV of the sum*

$$\textit{if } (\hat{\mathbf{P}}_A)_{\text{w}} = (\hat{\mathbf{P}}_B + \hat{\mathbf{P}}_C)_{\text{w}} \qquad \textit{then} \quad (\hat{\mathbf{P}}_A)_{\text{w}} = (\hat{\mathbf{P}}_B)_{\text{w}} + (\hat{\mathbf{P}}_C)_{\text{w}} \tag{2.1}$$

Proof. From linearity $\frac{\langle\Psi_{\text{fin}}|\hat{\mathbf{P}}_B+\hat{\mathbf{P}}_C|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{in}}\rangle} = \frac{\langle\Psi_{\text{fin}}|\hat{\mathbf{P}}_B|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{in}}\rangle} + \frac{\langle\Psi_{\text{fin}}|\hat{\mathbf{P}}_C|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{in}}\rangle}$. □

Theorem 3. *If a single IM of an observable $\hat{\mathbf{P}}_A$ is performed between the pre- and post-selection, then if the outcome is definite (e.g.* $\text{Prob}(\hat{\mathbf{P}}_A = 1) = 1$*) then the WV is equal to this eigenvalue (e.g.* $(\hat{\mathbf{P}}_A)_{\text{w}} = 1$*) [6].*

Proof. Given that $\hat{\mathbf{P}}_A = \sum_n a_n|\alpha_n\rangle\langle\alpha_n|$, if an eigenvalue, e.g. $\hat{\mathbf{P}}_A = a_n$, is obtained with certainty, then for $n \neq m, \hat{\mathbf{P}}_A \equiv |\alpha_m\rangle\langle\alpha_m| = 0$ because the probability of obtaining another eigenvalue by ABL is $\propto \langle\Psi_{\text{fin}}|\alpha_m\rangle\langle\alpha_m|\Psi_{\text{in}}\rangle = 0$. In this case, the weak value $(\hat{\mathbf{P}}_A)_{\text{w}} = (|\alpha_m\rangle\langle\alpha_m|)_{\text{w}} = \frac{\langle\Psi_{\text{fin}}|\alpha_m\rangle\langle\alpha_m|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{in}}\rangle} = 0$. In addition, $\sum_m \frac{\langle\Psi_{\text{fin}}|\alpha_m\rangle\langle\alpha_m|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{in}}\rangle} = 1$ because $\sum_m |\alpha_m\rangle\langle\alpha_m| = 1$. But since $\langle\Psi_{\text{fin}}|\alpha_m\rangle\langle\alpha_m|\Psi_{\text{in}}\rangle = 0$ for $n \neq m$, the only term left is n. Therefore, the weak value is 1, the same as the ideal value. □

This theorem also provides a direct link to the counter-factual statements (subsection 1.4) because all counter-factual statements which claim that something occurs with certainty, and which can actually be experimentally verified by *separate* IMs, continue to remain true when tested by WMs. However, given that WMs do not disturb each other, all these statements can be measured *simultaneously*.

Applying theorem 3 to the three boxes, we ascertain the following WVs with certainty

$$(\hat{\mathbf{P}}_A)_{\text{w}} = 1, \qquad (\hat{\mathbf{P}}_B)_{\text{w}} = 1, \qquad \hat{\mathbf{P}}_{\text{total}} = (\hat{\mathbf{P}}_A + \hat{\mathbf{P}}_B + \hat{\mathbf{P}}_C)_{\text{w}} = 1. \tag{2.2}$$

Using theorem 2, we also obtain

$$(\hat{\mathbf{P}}_C)_{\rm w} = \frac{\langle \Psi_{\rm fin}|\hat{\mathbf{P}}_{\rm total} - \hat{\mathbf{P}}_A - \hat{\mathbf{P}}_B|\Psi_{\rm in}\rangle}{\langle \Psi_{\rm fin}|\Psi_{\rm in}\rangle}$$
$$= (\hat{\mathbf{P}}_A + \hat{\mathbf{P}}_B + \hat{\mathbf{P}}_C)_{\rm w} - (\hat{\mathbf{P}}_A)_{\rm w} - (\hat{\mathbf{P}}_B)_{\rm w} = -1. \qquad (2.3)$$

This surprising theoretical prediction of TSQM has been verified experimentally using photons [34]. What interpretation should be given to $(\hat{\mathbf{P}}_C)_{\rm w} = -1$? Any WM which is sensitive to the projection operator $\hat{\mathbf{P}}_C$ will register the opposite effect from those cases in which the projection operator is positive, e.g. a WM of the amount of charge in box C in the intermediate time will yield a negative charge (assuming it is a positively charged particle). For numerous reasons [29, 50], we believe the most natural interpretation is: there are -1 particles in box C.

2.2. *Hidden variable theories, pre- and post-selection and value definiteness*

In the traditional constructions of the proofs of contextuality [23], the HVT had to first assign a definite, deterministic value to all possible observables of a system (an assumption called 'value definiteness'). Only after this step can noncontextuality be defined. A recent important development was demonstrated by Spekkens [48] who showed that noncontextuality could be defined for stochastic HVTs, thereby demonstrating that value definiteness and noncontextuality are distinct assumptions[4]. Notwithstanding these important results, we will consider the traditional treatment which requires value definiteness. The reason, as will be seen below, is that theorem 3 requires the analogue of value definiteness within ABL and PPS (i.e. in a similar spirit as is used within the subject of HVTs). We will use theorem 3 in a new notion called 'quantum contextuality' (defined below).

A 'logical PPS-paradox' is defined as [25] 'sets of projectors for which an ABL probability assignment violates the algebraic constraints, while every projector receives probability 0 or 1'. In [29], it was first pointed out and extensively discussed and later proven by Leifer and Spekkens [25] that there is a related proof of contextuality for 'logical PPS-paradoxes'.

Theorem 4. *'For every logical PPS-paradox with non-orthogonal pre- and post-selection projectors, there is an associated proof of the impossibility of an (measurement noncontextual) HVT that is obtained by considering all the measurements defined by the PPS paradox—the pre-selection measurement, the post-selection measurement and the alternative possible intermediate measurements—as alternative possible measurements at a single time' [25].*

Leifer and Spekkens [24] also argue that '...we did not show that a logical PPS-paradox is *itself* a proof of contextuality' because 'the measurements defined by the PPS-paradox—

[4] In [48], Spekkens extends BKS to operational theories, arbitrary preparations, transformations and unsharp measurements and to stochastic or indeterministic ontological models. Quantum states are given as probability distributions μ over HVTs λ, such that $\int_\Omega \mu(\lambda)\,{\rm d}\lambda = 1$, where Ω is the set of possible HVTs and measurements are characterized by idempotent indicator functions $\chi_j^{\rm M} : \Omega \to \{0, 1\}$, such that $\sum_j \chi_j^{\rm M}(\lambda) = 1$. The probability of obtaining an outcome j for a random variable X when a measurement M is performed on a system in μ then is $p_\mu(X_{\rm M} = j) = \int_\Omega \chi_j^{\rm M}(\lambda)\mu(\lambda)\,{\rm d}\lambda$. In addition, a transition from ω to some other HVT state λ could result from the measurement process [24], and this is modelled by a transition probability $D_j^{\rm M}(\lambda, \omega)$ such that $\int_\Omega D_j^{\rm M}(\lambda, \omega)\,{\rm d}\lambda = 1$. This approach to HVTs may then be applied [25] to ABL:

$$p_{\rm HVT}(X_{\rm M} = k|A_{\rm pre}, A_{\rm post}, {\rm M}) = \frac{\int_\Omega \chi_{\rm post}(\lambda)\Gamma_k^{\rm M}(\lambda, \omega)\mu_{\rm pre}(\omega){\rm d}\omega\,{\rm d}\lambda}{\int_\Omega \chi_{\rm post}(\lambda)(\Gamma_k^{\rm M}(\lambda, \omega) + \Gamma_{\neg k}^{\rm M}(\lambda, \omega))\mu_{\rm pre}(\omega)\,{\rm d}\omega\,{\rm d}\lambda}.$$

WMs and WVs can also be described in this language (the usual projective measurement typically utilized in quantum experiments is a special case of WMs [45]). In the HVT language of [25], logical PPS-paradox occurs if $p_{\rm HVT} = 1$ for several incompatible situations, e.g. if $\int_\Omega \chi_{\rm post}(\lambda)\Gamma_{\neg k}^{\rm M}(\lambda, \omega)\mu_{\rm pre}(\omega)\,{\rm d}\omega\,{\rm d}\lambda = 0$. Nevertheless [25] assume 'outcome determinism', i.e. 'the probability assigned to every projector for a particular ontic state is either 0 or 1'.

the pre-selection, the post-selection, and the alternative possible intermediate measurements,' while normally considered as 'temporal successors,' within the proof of contextuality, 'must be treated as counter-factual alternatives'. These counter-factual elements in the proof are valid only if we perform IMs *separately*; they do not hold if IMs are made *simultaneously*. This indeed is the essence of how counter-factual paradoxes are usually avoided and is also the same reason that BKS is generally thought not to be directly testable, i.e. BKS considers sets of non-commuting operators (which are considered in a counter-factual sense). Indeed, Leifer and Spekkens correctly argue that 'this distinction is critical, since an earlier measurement can cause a disturbance to the ontic state that is monitored by a later measurement' [24].

2.3. *Quantum contextuality*

However, WVs and WMs (instead of the IMs used in theorem 4) give a different meaning to 'alternative possible measurements at a single time'. Although these alternatives are usually regarded as counter-factual, they will all be true simultaneously with WVs which can be probed empirically—to some extent—by WMs because theorem 3 provides a direct link with all the counter-factual alternatives in theorem 4. Theorem 3 says that all counter-factual statements which claim that something occurs with certainty, and which can actually be experimentally verified by *separate* ideal experiments, continue to remain true *simultaneously* as WVs. We will first emphasize the purely theoretical aspect of WVs as distinct from attempts to measure them with WMs. Although WVs require an ensemble in order to be probed empirically with WMs, we can consider WVs by themselves as a general, non-statistical and robust mathematical property of every individual PPS system (see appendix B).

As a result of this new empirical access through WMs, rather than focusing on HVTs, we will be interested in the relevance of contextuality for QM itself. As a consequence, we will define a new concept called 'quantum contextuality' which is directly applicable to QM. Although 'quantum contextuality' is not based on HVTs, it does follow the spirit of their definition.

> *Quantum contextuality.* For any initial quantum state which exhibits a breakdown of noncontextuality in the associated HVT for a certain set of operators (i.e. for which ABL assigns definite values of $0/1$), one can find at least one post-selected state which will show how the function composition rule (i.e. sum and product rules) is violated.

This definition focuses our inquiries into the specific conditions under which the function composition rule (1.1) is violated. That is, 'quantum contextuality' refers to those situations where the function composition is satisfied for one boundary condition (i.e. the pre- *or* post-selection) but not for both the pre- *and* the post-selection.

2.4. *Logical PPS-paradoxes imply quantum contextuality through weak values*

We are now able to prove the following.

Theorem 5. *Logical PPS-paradoxes imply 'quantum contextuality' through WVs.*

Proof. The proof follows [25] but circumvents the counter-factual status of measurements that are temporal successors by using theorem 3 which allows us to state that all counter-factual statements which maintain the occurrences of an outcome with certainty will all be true simultaneously when they are measured weakly. Theorem 3 is applicable to the precise elements utilized in the contextuality proof [25]. In addition, given that WVs are by definition independent of the type of WM and given equation (1.10), and since WVs can violate the algebraic conditions (the product rule) required for BKS and noncontextual HVTs [24], we have now proven that logical PPS-paradoxes which assign definite probabilities (of 0 or 1) via ABL, are in fact proofs of contextuality if all the 'alternative possible intermediate measurements' are performed weakly. $\square$

In addition, for any logical PPS-paradox, we can always find a post-selection which can empirically manifest non-classical WVs as implied in the definition of quantum contextuality. To see this, consider that for every two commuting projection operators, we can always find a unitary operation to obtain a representation $\Pi_1 \equiv |1\rangle\langle 1|$ and $\Pi_2 \equiv |2\rangle\langle 2|$. By way of example, in three dimensions, the most general pre-selected vector can always be written in the form

$$|\Psi_{\text{in}}\rangle = a|1\rangle + b|2\rangle + c|3\rangle. \tag{2.4}$$

Therefore, there is always a post-selected vector

$$|\Psi_{\text{fin}}\rangle = \frac{1}{a}|1\rangle + \frac{1}{b}|2\rangle - \frac{1}{c}|3\rangle \tag{2.5}$$

such that $(\Pi_1)_w = 1$ and $(\Pi_2)_w = 1$. Using $(\Pi_1\Pi_2)_w = 0$ and $\Pi_1 + \Pi_2 + \Pi_3 = 1$, we obtain the non-classical WV $(\Pi_3)_w = -1$. Similarly, in four dimensions, the most general pre-selected vector is $|\Psi_{\text{in}}\rangle = a|1\rangle + b|2\rangle + c|3\rangle + d|4\rangle$ and the post-selected vector $|\Psi_{\text{fin}}\rangle = \frac{1}{a}|1\rangle + \frac{1}{b}|2\rangle + \frac{1}{c}|3\rangle - \frac{2}{d}|4\rangle$ will reveal the non-classical WV. This continues in the same way for any dimension.

We will argue (section 5) that these considerations mitigate the attempt to explain these PPS 'paradoxes' as a result of disturbance (since there is no disturbance with WMs) and have strengthened [25] the connection between these 'paradoxes' and contextuality. The following two sections will explore worked examples of these new considerations in higher dimensions.

3. Pre- and post-selection and contextuality in four dimensions

In the three-box paradox, the product of observables was always definite, i.e. $\hat{\mathbf{P}}_A\hat{\mathbf{P}}_B = 0$ and the proof of contextuality was state dependent. In this section, we consider a slightly different situation (4D BKS nonets) in which the product of observables can give two different values. Except for this difference, this 4D BKS nonet example is similar to the three-box paradox in that we shall also analyse them in terms of PPSs thereby revealing surprising predictions for IMs and WMs and will demonstrate the identical issues of diagonal-PPS measurements, violation of the product rule, contextuality and WMs which cannot be explained by a noncontextual HVT.

3.1. Review of Mermin's 4D BKS theorem

We consider Mermin's version of BKS with a set of nine observables. It is intuitive [27] to represent all the 'functional relationships between mutually commuting subsets of the observables,' i.e. $V_{\bar{\psi}}(F\{\hat{A}\}) = F\{V_{\bar{\psi}}(\hat{A})\}$, by drawing them in figure 4 and arranging them so that all the observables in each row (and column) commute with all the other observables in the same row (or column).

Individually, each of the nine observables depicted in figure 4 has eigenvalues ± 1. In addition, equation (1.1) requires that the value assigned to the product of all three observables in any row or column must obey the same identities that the observables themselves satisfy, i.e. the product of the values assigned to the observables in each oval yields a result of $+1$ except in the last column which gives -1. (The value assignments are given by

$V_{\vec{\psi}}(\hat{\sigma}_x^1) = \langle \hat{\sigma}_x^1 \otimes I^2 \rangle$, $V_{\vec{\psi}}(\hat{\sigma}_x^2) = \langle I^1 \otimes \hat{\sigma}_x^2 \rangle \cdots V_{\vec{\psi}}(\hat{\sigma}_x^1) = \langle \hat{\sigma}_z^1 \otimes \hat{\sigma}_z^2 \rangle$). Computing column 3 of figure 4,

$$\begin{aligned}\{\hat{\sigma}_x^1\hat{\sigma}_x^2\}\{\hat{\sigma}_y^1\hat{\sigma}_y^2\}\{\hat{\sigma}_z^1\hat{\sigma}_z^2\} &= \hat{\sigma}_x^1 \underbrace{\hat{\sigma}_x^2\hat{\sigma}_y^1}_{\text{commute so}\hookrightarrow} \hat{\sigma}_y^2\hat{\sigma}_z^1\hat{\sigma}_z^2 = \underbrace{\hat{\sigma}_x^1\hat{\sigma}_y^1}_{=\mathrm{i}\hat{\sigma}_z^1}\underbrace{\hat{\sigma}_x^2\hat{\sigma}_y^2}_{=\mathrm{i}\hat{\sigma}_z^2}\hat{\sigma}_z^1\hat{\sigma}_z^2 \\ &= \mathrm{i}\hat{\sigma}_z^1 \underbrace{\mathrm{i}\hat{\sigma}_z^2\hat{\sigma}_z^1}_{\text{commute so}\hookrightarrow} \hat{\sigma}_z^2 = \mathrm{i}\hat{\sigma}_z^1\mathrm{i}\hat{\sigma}_z^1\hat{\sigma}_z^2\hat{\sigma}_z^2 = -1. \end{aligned} \tag{3.1}$$

Computing the product of the observables in the third row, i.e.,

$$\begin{aligned}\{\hat{\sigma}_x^1\hat{\sigma}_y^2\}\{\hat{\sigma}_x^2\hat{\sigma}_y^1\}\{\hat{\sigma}_z^1\hat{\sigma}_z^2\} &= \hat{\sigma}_x^1 \underbrace{\hat{\sigma}_y^2\hat{\sigma}_x^2}_{=-\mathrm{i}\hat{\sigma}_z^2} \hat{\sigma}_y^1\{\hat{\sigma}_z^1\hat{\sigma}_z^2\} = \underbrace{\hat{\sigma}_x^1\hat{\sigma}_y^1}_{=\mathrm{i}\hat{\sigma}_z^1}\underbrace{\{-\mathrm{i}\hat{\sigma}_z^2\}\{\hat{\sigma}_z^1}_{\text{commute so}\hookrightarrow}\hat{\sigma}_z^2\} \\ &= \underbrace{\mathrm{i}\hat{\sigma}_z^1\hat{\sigma}_z^1}_{=\mathrm{i}}\underbrace{\{-\mathrm{i}\hat{\sigma}_z^2\}\{\hat{\sigma}_z^2\}}_{=-\mathrm{i}} = +1. \end{aligned} \tag{3.2}$$

If the product rule is applied to the value assignments made in the rows, then

$$\begin{aligned} V_{\vec{\psi}}(\hat{\sigma}_x^1)V_{\vec{\psi}}(\hat{\sigma}_x^2)V_{\vec{\psi}}(\hat{\sigma}_x^1\hat{\sigma}_x^2) &= V_{\vec{\psi}}(\hat{\sigma}_y^2)V_{\vec{\psi}}(\hat{\sigma}_y^1)V_{\vec{\psi}}(\hat{\sigma}_y^1\hat{\sigma}_y^2) \\ &= V_{\vec{\psi}}(\hat{\sigma}_x^1\hat{\sigma}_y^2)V_{\vec{\psi}}(\hat{\sigma}_x^2\hat{\sigma}_y^1)V_{\vec{\psi}}(\hat{\sigma}_z^1\hat{\sigma}_z^2) = +1 \end{aligned} \tag{3.3}$$

while the column identities require

$$\begin{aligned} &V_{\vec{\psi}}(\hat{\sigma}_x^1)V_{\vec{\psi}}(\hat{\sigma}_y^2)V_{\vec{\psi}}(\hat{\sigma}_x^1\hat{\sigma}_y^2) = V_{\vec{\psi}}(\hat{\sigma}_x^2)V_{\vec{\psi}}(\hat{\sigma}_y^1)V_{\vec{\psi}}(\hat{\sigma}_x^2\hat{\sigma}_y^1) = +1 \\ &V_{\vec{\psi}}(\hat{\sigma}_x^1\hat{\sigma}_x^2)V_{\vec{\psi}}(\hat{\sigma}_y^1\hat{\sigma}_y^2)V_{\vec{\psi}}(\hat{\sigma}_z^1\hat{\sigma}_z^2) = -1 \end{aligned} \tag{3.4}$$

However, it is easy to show that the nine numbers $V_{\vec{\psi}}$ cannot satisfy all six constraints: multiplying all nine values together gives two different results, a +1 when it is done row by row and a −1 when it is done column by column

$$V_{\vec{\psi}}(\hat{\sigma}_x^1)V_{\vec{\psi}}(\hat{\sigma}_x^2)V_{\vec{\psi}}(\hat{\sigma}_x^1\hat{\sigma}_x^2)\cdots V_{\vec{\psi}}(\hat{\sigma}_z^1\hat{\sigma}_z^2) = +1 \tag{3.5}$$

$$V_{\vec{\psi}}(\hat{\sigma}_x^1)V_{\vec{\psi}}(\hat{\sigma}_y^2)V_{\vec{\psi}}(\hat{\sigma}_x^1\hat{\sigma}_y^2)\cdots V_{\vec{\psi}}(\hat{\sigma}_z^1\hat{\sigma}_z^2) = -1. \tag{3.6}$$

There obviously is no consistent solution to equations (3.6) and (3.5) since they contain the same set of numbers, simply ordered differently. Therefore the values assigned to the observables cannot obey the same identities that the observables themselves obey, $V_{\vec{\psi}}(F\{\hat{A}\}) \neq F\{V_{\vec{\psi}}(\hat{A})\}$, and an HVT would have to assign values to observables in a way that depended on the choice of which of two mutually commuting sets of observables that were also chosen to measure, i.e. the values assigned are contextual.

3.2. ABL, VAA and BKS nonets

Following Vaidman, Albert, and Aharonov (VAA) [4], Mermin showed how to assign a definite value to a single measurement of any one of the nine observables of a BKS nonet [26]. He then generalized this to a definite assignment to any one of 16 observables from the sets of nonets and showed that this assignment cannot be done if one attempts to measure (or ascertain) two or more of the observables belonging to the nonets. He left open the question as to the physical reason for this, stating 'I find this intriguing' [26]. To address this, we present a physical reason to demonstrate why the VAA scheme cannot be applied to two or more measurements by showing that the two measurements interfere with each other given the necessary PPSs. This can be seen to be a consequence of TSQM: some assignments of eigenvalues to operators are based on just one of the two vectors (i.e. either the pre- *or* the post-selected vector) while

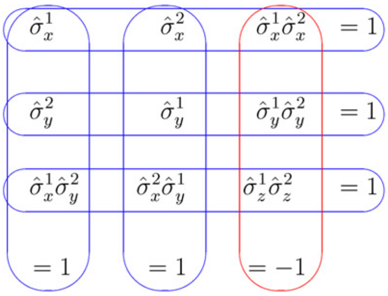

Figure 4. 4D BKS example.

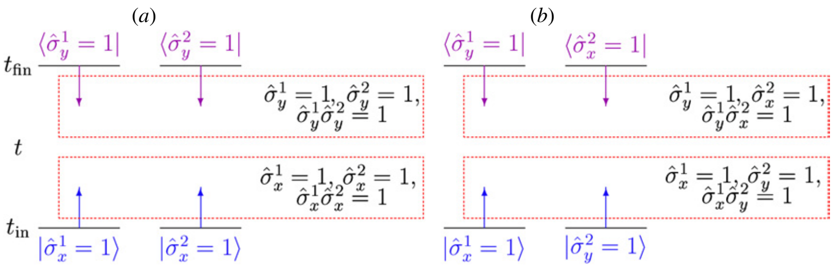

Figure 5. PPS states for Mermin example.

some assignments of eigenvalues are based on both vectors (i.e. both the pre-selected *and* the post-selected vectors—what we call diagonal-PPS). In this picture, it is the utilization of more than one PPS and the subsequent interference between them that explains the violation of the product rule and thus the physical source of the 'contextuality'. When assignments are not made in the diagonal-PPS sense, then sets of commuting observables which are determined entirely by just one vector *satisfy* the BKS function condition $V_{\vec{\psi}}(F\{\hat{A}\}) = F\{V_{\vec{\psi}}(\hat{A})\}$. Sets of commuting observables which are assigned values in the diagonal-PPS sense by using information from both vectors *do not* satisfy the BKS function condition because they violate the product rule, and can disturb each other.

3.2.1. Ascertaining the results of any one of the nine observables. We begin our analysis by considering specific examples of PPS configurations. We then utilize ABL [2] and show that choosing different post-selections changes the triplet of observables that violate the product rule. Consider first a pre-selection of $\hat{\sigma}_x^1 = 1$ and $\hat{\sigma}_x^2 = 1$ and a post-selection of $\hat{\sigma}_y^1 = 1$ and $\hat{\sigma}_y^2 = 1$ (see figure 5(*a*)). In this case, it is easy to see that we can ascertain with certainty any one of the following values: $\hat{\sigma}_x^1 = 1, \hat{\sigma}_x^2 = 1, \hat{\sigma}_y^1 = 1$ or $\hat{\sigma}_y^2 = 1$. We also know that we will obtain definite values of +1 if we measure any one of the following products of observables: $\hat{\sigma}_x^1\hat{\sigma}_x^2 = 1$ and $\hat{\sigma}_y^1\hat{\sigma}_y^2 = 1$, so we must also obtain $\hat{\sigma}_x^1\hat{\sigma}_x^2\hat{\sigma}_y^1\hat{\sigma}_y^2 = +1$:

$$\left\langle\hat{\sigma}_x^1\hat{\sigma}_x^2\hat{\sigma}_y^1\hat{\sigma}_y^2\right\rangle = \left\langle\hat{\sigma}_y^1 = 1\right|\left\langle\hat{\sigma}_y^2 = 1\right|\hat{\sigma}_x^1\hat{\sigma}_x^2\hat{\sigma}_y^1\hat{\sigma}_y^2\left\rangle\right|\hat{\sigma}_x^1 = 1\left.\right\rangle\left|\hat{\sigma}_x^2 = 1\right\rangle = +1. \tag{3.7}$$

In addition, we obtain the same results if we switch the sequence of $\hat{\sigma}_x^1\hat{\sigma}_x^2$ and $\hat{\sigma}_y^1\hat{\sigma}_y^2$ because they commute. Any *one* of the other observables in figure 4 (i.e. $\hat{\sigma}_x^1\hat{\sigma}_y^2$ and $\hat{\sigma}_x^2\hat{\sigma}_y^1$) can also be ascertained with certainty given this PPS. Finally, from the product of the three observables in column 3 $\left(\{\hat{\sigma}_x^1\hat{\sigma}_x^2\}\{\hat{\sigma}_y^1\hat{\sigma}_y^2\}\{\hat{\sigma}_z^1\hat{\sigma}_z^2\} = -1 \text{ and } \{\hat{\sigma}_x^1\hat{\sigma}_x^2\}\{\hat{\sigma}_y^1\hat{\sigma}_y^2\} = +1\right)$, we can deduce that

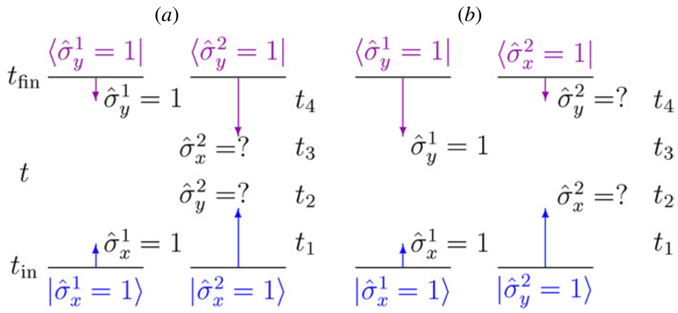

Figure 6. Time sequence of PPS measurements for Mermin example.

$\hat\sigma_z^1\hat\sigma_z^2 = -1$. Similar statements can be made for other PPSs. For example, consider the pre-selection $\hat\sigma_x^1 = 1$ and $\hat\sigma_y^2 = 1$ and a post-selection of $\hat\sigma_y^1 = 1$ and $\hat\sigma_x^2 = 1$ (see figure 5(*b*)). In this case, the measurement $\hat\sigma_x^1\hat\sigma_y^2\hat\sigma_x^2\hat\sigma_y^1 = +1$, and we can deduce from the third row that $\left\{\hat\sigma_z^1\hat\sigma_z^2\right\} = +1$.

Thus, given just one PPS, any single observable can be assigned a definite value, even though $\hat\sigma_z^1\hat\sigma_z^2$ is assigned different values in different PPS. It is precisely because of this connection between particular PPSs and different values for $\hat\sigma_z^1\hat\sigma_z^2$ that the issue of contextuality arises when we consider products of these observables.

3.2.2. Ascertaining the results of products of the nine observables:. In this section, we ask how many of the *products* of the nine observables in figure 4 can be ascertained together with certainty. For example, as stated in the previous section, the outcome for the product of the first two observables in column 3 of figure 4 with the PPS of figure 5(*a*) is $\sigma_x^1\sigma_x^2\sigma_y^1\sigma_y^2 = +1$. However, if we measure the operators corresponding to the first two observables of row 3 in figure 4, i.e. $\hat\sigma_x^1\hat\sigma_y^2\hat\sigma_x^2\hat\sigma_y^1$, given this particular PPS shown in figure 5(*a*), then the sequence of measurements interferes with each other (as represented by the slanted ovals in figure 7(*a*)). To see this, consider that $\hat\sigma_x^1\hat\sigma_y^2\hat\sigma_x^2\hat\sigma_y^1$ corresponds to the sequence of measurements represented in figure 8(*a*). While the pre-selection of particle 2 is $\hat\sigma_x^2 = 1$ at $t_{\rm in}$, the next measurement after the pre-selection at t_2 is for $\hat\sigma_y^2$ and only *after* that a measurement of $\hat\sigma_x^2$ is performed at t_3. Thus, there is no guarantee that the $\hat\sigma_x^2$ measurement at t_3 will give the same value as the pre-selected state of $\hat\sigma_x^2 = 1$ or that the $\hat\sigma_y^2$ measurement will give the same value as the post-selected state of $\hat\sigma_y^2 = 1$. In TSQM, this is due to the disturbance of the 2-vector boundary conditions which is created by the IM: the initial pre-selected vector $\hat\sigma_x^2 = 1$ from $t_{\rm in}$ is 'destroyed' when the $\hat\sigma_y^2$ measurement at time t_2 is performed and therefore cannot inform the later $\hat\sigma_x^2$ measurement at time t_3. In other words, with the particular PPS given in figures 5(*a*) and 6(*a*), the operator, $\hat\sigma_x^1\hat\sigma_y^2\hat\sigma_x^2\hat\sigma_y^1$ depends on information from both the pre-selected vector $\hat\sigma_x^1 = 1, \hat\sigma_x^2 = 1$ and the post-selected vector $\hat\sigma_y^1 = 1, \hat\sigma_y^2 = 1$ in a diagonal-PPS sense. We call this diagonal-PPS because a line connecting $\hat\sigma_x^1(t_1)$ with $\hat\sigma_x^2(t_3)$ will be diagonal or will cross the line connecting $\hat\sigma_y^2(t_2)$ with $\hat\sigma_y^1(t_4)$, where $t_{\rm in} < t_1 < t_2 \cdots < t_{\rm fin}$ (see figure 7(*a*)).

These results can also be seen in an actual measurement situation, we consider interaction Hamiltonians with coupling terms $\hat\sigma_x^1\delta(t-t_1)$, $\hat\sigma_x^2\delta(t-t_2)$, $\hat\sigma_y^2\delta(t-t_2)$ and $\hat\sigma_y^1\delta(t-t_1)$:

$$\begin{aligned}\langle\hat\sigma_y^1 = 1|\langle\hat\sigma_y^2 = 1|\, \mathrm{e}^{\mathrm{i}\hat Q_1\hat\sigma_x^1\hat\sigma_x^2}\,\mathrm{e}^{\mathrm{i}\hat Q_2\hat\sigma_y^1\hat\sigma_y^2}|\hat\sigma_x^1 = 1\rangle|\hat\sigma_x^2 = 1\rangle \\ = \langle\hat\sigma_y^1 = 1|\langle\hat\sigma_y^2 = 1|\, \mathrm{e}^{\mathrm{i}\hat Q_2\hat\sigma_y^1\hat\sigma_y^2}\,\mathrm{e}^{\mathrm{i}\hat Q_1\hat\sigma_x^1\hat\sigma_x^2}|\hat\sigma_x^1 = 1\rangle|\hat\sigma_x^2 = 1\rangle = 1.\end{aligned} \tag{3.8}$$

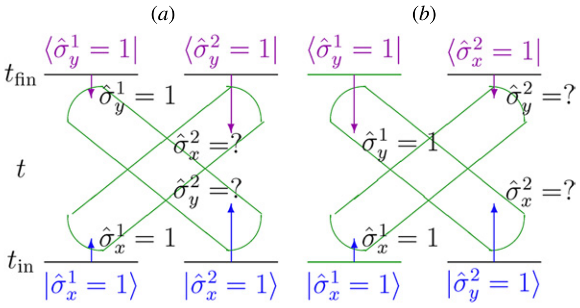

Figure 7. (*a*) Measurement of $\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1$ is diagonal. (*b*) Measurement of $\hat{\sigma}_x^1\hat{\sigma}_x^2\hat{\sigma}_y^1\hat{\sigma}_y^2$ is diagonal.

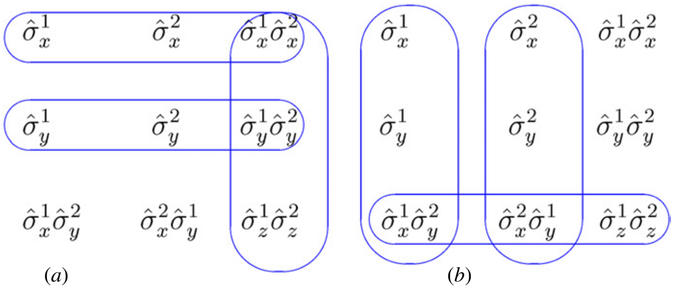

Figure 8. Products of observables that are not disturbed. (*a*) Given the PPS of figure 2(*a*). (*b*) Given the PPS of figure 2(b).

Since $e^{i\hat{Q}_1\hat{\sigma}_x^1\hat{\sigma}_x^2}$ commutes with $e^{i\hat{Q}_2\hat{\sigma}_y^1\hat{\sigma}_y^2}$, they can be interchanged and thus the same outcome of +1 is obtained. However, for the other observables $\langle\hat{\sigma}_y^1=1|\langle\hat{\sigma}_y^2=1|\,e^{i\hat{Q}_1\hat{\sigma}_x^1\hat{\sigma}_y^2}\,e^{i\hat{Q}_2\hat{\sigma}_y^1\hat{\sigma}_x^2}|\hat{\sigma}_x^1=1\rangle$ $|\hat{\sigma}_x^2=1\rangle$ the opposite eigenvalue is obtained (even though they commute) i.e. $\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1$ will give an outcome of -1 given the PPS of figure 6(*a*) even though separately $\hat{\sigma}_x^1\hat{\sigma}_y^2=+1$ and $\hat{\sigma}_x^2\hat{\sigma}_y^1=+1$. This is thus a violation of the product rule (see figure 7(*a*)). This diagonal-PPS phenomenon can be generalized to functions that are polynomials of products of observables with the proper ordering (i.e. no mixing or sandwiching).

To summarize this sub-section, given the PPS of figure 7(*a*), the subset of observables circled in figure 8(*a*) (and the products of those circled observables) can be assigned eigenvalues in a way that satisfies the function relation requirement equation (1.1). But, the product of the other observables (e.g. $\hat{\sigma}_x^1\hat{\sigma}_y^2$ and $\hat{\sigma}_x^2\hat{\sigma}_y^1$) can only be ascertained (given this particular PPS) using information from *both* the pre- and post-selected vector in a diagonal sense (see figure 7(*a*)), and will thus violate the product rule. With the PPS of figure 7(*b*), the subset in figure 8(*b*) (and the relevant products of observables) can be assigned eigenvalues in a way that satisfies the function relation requirement equation (1.1). But, the product of the other observables $\hat{\sigma}_x^1\hat{\sigma}_x^2$ and $\hat{\sigma}_y^1\hat{\sigma}_y^2$, i.e. $\hat{\sigma}_x^1\hat{\sigma}_x^2\hat{\sigma}_y^1\hat{\sigma}_y^2$ violates the function rule equation (1.1).

3.2.3. Ascertaining the results of any one of 16 observables through the generalized state. As was explained in subsubsection 3.2.1, definite results for any *one* of the intermediate measurements can be obtained for several different PPSs which were complete measurements (and thus describable by a wavefunction). In addition, some triplets (products of observables)

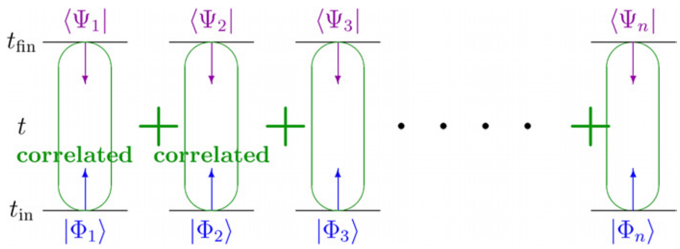

Figure 9. Generalized state: superpositions of 2-vectors given by equation (3.11).

can also be ascertained (see figure 8) given a particular pre- or post-selection. However, for the most general setup considered in this section, no two products of observables can be ascertained. The general setup considers superpositions of these PPSs. In fact, this is required in order to ascertain any one of the 16 observables in Mermin's successful generalization of VAA (it is also required to ascertain any one of the nine for general PPSs). Following VAA [4], Mermin showed that the way that any one of the 16 values of the 7 BKS nonets (e.g. figure 4 is one of them, the others given by $\sigma_\nu\sigma_\mu$) can be ascertained with certainty is by entangling the two-particle system representing the 4D BKS nonet (represented by the pre- and post-selection of figures 5(*a*) and (*b*), etc, and labelled as $|\mathrm{i}\rangle$) with another system, i.e. an ancilla (represented here by $|\Psi_i\rangle$ and $|\Phi_i\rangle$). The nonet is prepared at t_{in} by correlating it with a set of states of an ancilla:

$$|A\rangle = \sum_i \frac{1}{\sqrt{N}} |\Phi_i\rangle|\mathrm{i}\rangle. \tag{3.9}$$

Then the ancilla is 'guarded' so there are no interactions with the ancilla during the time $(t_{\mathrm{in}}, t_{\mathrm{fin}})$. At t_{fin} we post select on the particle and ancilla and obtain the state

$$|\Omega\rangle = \frac{1}{\sqrt{N}} \sum_i |\Psi_i\rangle|\mathrm{i}\rangle. \tag{3.10}$$

If we are successful in obtaining this state for the post-selection, then the state of the system is described in the intermediate time by the entangled state (see figure 9) [6, 9],

$$\Psi = \sum_i \alpha_i \langle\Psi_i||\Phi_i\rangle. \tag{3.11}$$

For general PPSs, we use multiple sets of boundary conditions given by figures 5(*a*) and (*b*), etc, to get an entangled state represented by figure 10 (where for simplicity we have taken the states of the ancilla to be an orthonormal set, $\mathbf{f}^\mu$). Mermin then presented an elegant method to determine the states of the ancilla necessary to produce the effect: he selected a definite representation for the two-particle spins, performed a projection, $\hat{\mathbf{P}}$, onto the subspace given by the nonet and solved under rotation for the state of the ancilla. That is, $\langle\Omega|\hat{\mathbf{P}}|A\rangle = 0$ for 'all but a single one of the projections associated with . . . ' [26] the observable of the nonet that is to be ascertained with certainty.

As Mermin proved, all the four components of this generalized state are necessary (i.e. we only determine a Bell state at t_{fin} on the ancilla rather than making a projection onto any given component $\mathbf{f}^\mu$) to ascertain a definite answer to any one of the individual observables.

3.2.4. A physical reason for restrictions on these assignments. We have suggested a physical reason based on TSQM and PPS for the two different values for $\hat{\sigma}_z^1\hat{\sigma}_z^2$. This points to a physical

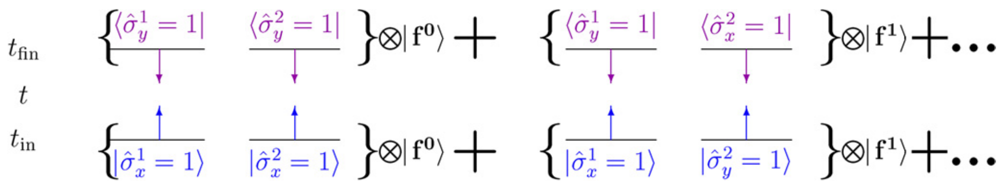

Figure 10. Generalized state for BKS nonets.

reason why no two measurements can be ascertained with certainty in the intermediate[5] time: all sets of boundary conditions are needed (those corresponding to both figures 5(*a*) and (*b*), etc) in order to ascertain with certainty the value of any one of the 16 observables, as represented by figure 10. However, when the first observable is ascertained, then it will depend on both the pre- and the post-selection measurements (i.e. it will be diagonal-PPS) in two of the four components of the generalized state (see subsubsection 3.2.2) and will collapse the entire configuration onto a subset of the PPSs, thereby disturbing the terms of the generalized state. Given any pair of measurements, there will always be a diagonal-PPS situation when all the four components of the generalized state are considered. This can be seen by comparing figures 8(*a*) and (*b*) and noting that any two observables will not be circled in both. Therefore, since we cannot be sure that the entire setup (see figure 10) is not disturbed, we cannot ascertain with certainty the outcome for any one of the 16 observables for the second measurement. Furthermore, this arrangement is the maximal correlation that can be performed (i.e. the 4D state of two spins can be maximally correlated to another 4D system as performed here), and thus we cannot create an even more sophisticated situation with additional ancillas. We have thus given a physical picture for Mermin's 'intriguing' question: there will always be a diagonal situation for any two observables.

3.3. Non-classical weak values for the 4D BKS nonets

We can now clarify Mermin's statement: 'Alice's other two 'results' have nothing to do with any properties of the particle or the results of any measurement actually performed' [26]. While it is certainly true that these 'other results' cannot be ascertained simultaneously in terms of an IM (as was demonstrated in subsubsection 3.2.4 and by Mermin), they *can* be measured simultaneously through WMs. The route to an easy calculation of WVs can be established from Mermin's description of VAA's accomplishment: 'Alice's list gives the observed result for the measurement Bob actually made and had he measured anything else it would have given the result he observed' [26]. This provides a direct route to WMs through theorem 3: WMs will produce the identical result as predicted for the IM since the IM results are definite. Thus, the other 'results' are related to properties of the particle and can be simultaneously measured.

We can also obtain non-classical results in this example (similar to the three-box paradox), by first rewriting the observables of figure 4 in terms of three-spin components for two 'virtual' particles: for the first particle, $\hat{\mathbf{S}}^{\mathbf{1}}_{\mathbf{3}} \equiv \hat{\sigma}_x^1\hat{\sigma}_y^2$, $\hat{\mathbf{S}}^{\mathbf{1}}_{\mathbf{2}} \equiv \hat{\sigma}_x^1\hat{\sigma}_z^2$, and $\hat{\mathbf{S}}^{\mathbf{1}}_{\mathbf{1}} \equiv \hat{\sigma}_x^2$; for particle 2, components which commute with particle 1 are $\hat{\mathbf{S}}^{\mathbf{2}}_{\mathbf{3}} \equiv \hat{\sigma}_x^2\hat{\sigma}_y^1$, $\hat{\mathbf{S}}^{\mathbf{2}}_{\mathbf{2}} \equiv \hat{\sigma}_z^1\hat{\sigma}_x^2$, and $\hat{\mathbf{S}}^{\mathbf{2}}_{\mathbf{1}} \equiv \hat{\sigma}_x^1$. We can observe a non-classical WV by noting that $\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1 = -1$ given the PPS of figure 6(*a*) even though separately $\hat{\sigma}_x^1\hat{\sigma}_y^2 = +1$ and $\hat{\sigma}_x^2\hat{\sigma}_y^1 = +1$, i.e. a violation of the product rule and thus a diagonal

[5] If we were considering a single PPS, as discussed in subsubsections 3.2.1 and 3.2.2, then *some* pairs of products of observables can be ascertained with certainty, but not any two pairs. No two products can be ascertained in the case of 16 observables and some two pairs of products cannot be ascertained for the nine observables.

situation (see figure 7(*a*)). Thus WMs must yield the same outcomes, i.e. $\hat{\mathbf{N}}(\hat{\mathbf{S}}_3^1)_w = +1$ and $\hat{\mathbf{N}}(\hat{\mathbf{S}}_3^2)_w = +1$ but $\hat{\mathbf{N}}(\hat{\mathbf{S}}_3^1\hat{\mathbf{S}}_3^2)_w = -1$, a non-classical result. To analyse these results, we define the following pair occupation operators:

$\hat{\mathbf{N}}_{++}$ the projector on the state $\hat{\mathbf{S}}_3^1 = 1$ and $\hat{\mathbf{S}}_3^2 = 1$
$\hat{\mathbf{N}}_{+-}$ the projector on the state $\hat{\mathbf{S}}_3^1 = 1$ and $\hat{\mathbf{S}}_3^2 = -1$
$\hat{\mathbf{N}}_{-+}$ the projector on the state $\hat{\mathbf{S}}_3^1 = -1$ and $\hat{\mathbf{S}}_3^2 = 1$
$\hat{\mathbf{N}}_{--}$ the projector on the state $\hat{\mathbf{S}}_3^1 = -1$ and $\hat{\mathbf{S}}_3^2 = -1$.

We can relate these measurements to the three box example of subsection 1.3, but in this case we have two boxes and two particles: $\hat{\mathbf{N}}_{++}$ means the number of times that particles 1 and 2 are in the first box, $\hat{\mathbf{N}}_{+-}$ means the number of times that particle 1 is in the first box and particle 2 is in the second box, etc. The WV of the projection operator $(1 - \hat{\mathbf{S}}_3^1)(1 - \hat{\mathbf{S}}_3^2)$ is $-\frac{1}{2}$, a non-classical result. From theorems 2 and 3, we can deduce the following: the different ways to obtain $\hat{\mathbf{N}}(\hat{\mathbf{S}}_3^2) = +1$ are given by $\hat{\mathbf{N}}_{++}$ (i.e. $\hat{\mathbf{S}}_3^1 = 1$ and $\hat{\mathbf{S}}_3^2 = 1$) and $\hat{\mathbf{N}}_{-+}$ (i.e. $\hat{\mathbf{S}}_3^1 = -1$ and $\hat{\mathbf{S}}_3^2 = 1$) and therefore we can deduce that

$$\hat{\mathbf{N}}_{*+} \equiv \hat{\mathbf{N}}_{++} + \hat{\mathbf{N}}_{-+} = 1. \tag{3.12}$$

In terms of the box analogy, this is how many ways that particle 2 can be found in box 1. Also $\hat{\mathbf{N}}(\hat{\mathbf{S}}_3^1) = +1$, and thus $\hat{\mathbf{N}}(\hat{\mathbf{S}}_3^1) \neq -1$ (i.e. how many ways can particle 1 be found in box 1). This is characterized by

$$\hat{\mathbf{N}}_{+*} \equiv \hat{\mathbf{N}}_{++} + \hat{\mathbf{N}}_{+-} = 1. \tag{3.13}$$

From $\hat{\mathbf{S}}_3^1\hat{\mathbf{S}}_3^2 = -1$ (which again means that both particles cannot be found in the same box), it cannot be that $\hat{\mathbf{S}}_3^2\hat{\mathbf{S}}_3^1 = +1$ and thus $\hat{\mathbf{S}}_3^2$ and $\hat{\mathbf{S}}_3^1$ must be opposite in sign. This can be characterized by

$$\hat{\mathbf{N}}_{++} + \hat{\mathbf{N}}_{--} = 0. \tag{3.14}$$

Furthermore, since equation (3.12) equals equation (3.13), we can deduce

$$\hat{\mathbf{N}}_{+-} = \hat{\mathbf{N}}_{-+}. \tag{3.15}$$

Subtracting equation (3.14) from the following identity:

$$\hat{\mathbf{N}}_{++} + \hat{\mathbf{N}}_{--} + \hat{\mathbf{N}}_{+-} + \hat{\mathbf{N}}_{-+} = +1, \tag{3.16}$$

we obtain

$$\hat{\mathbf{N}}_{+-} + \hat{\mathbf{N}}_{-+} = 1. \tag{3.17}$$

Equation (3.17) implies that the two particles are never in the same box. From equations (3.15) and (3.17), we can deduce

$$\hat{\mathbf{N}}_{-+} = \hat{\mathbf{N}}_{+-} = \tfrac{1}{2}. \tag{3.18}$$

Substituting this value into equation (3.13), we can deduce that

$$\hat{\mathbf{N}}_{++} = \tfrac{1}{2}. \tag{3.19}$$

Finally, substituting this into equation (3.14), we can deduce

$$\hat{\mathbf{N}}_{--} = -\tfrac{1}{2}. \tag{3.20}$$

As shown in [11], all these statements can be measured simultaneously through WMs and will yield

$$(\hat{\mathbf{N}}_{-+})_w = (\hat{\mathbf{N}}_{+-})_w = (\hat{\mathbf{N}}_{++})_w = \tfrac{1}{2}, \tag{3.21}$$

while

$$(\hat{\mathbf{N}}_{--})_{\mathrm{w}} = -\tfrac{1}{2}. \tag{3.22}$$

In other words, if a WM is performed on the number of times that particle 1 is in the second box and particle 2 is in the second box, then the result is the non-classical result $-\frac{1}{2}$. Thus, the way that the two seemingly contradictory statements $\hat{\sigma}_z^1\hat{\sigma}_z^2 = \pm 1$ can weakly 'peacefully co-exist' is that a WV goes outside the spectrum of possible eigenvalues, i.e. equation (3.22). In summary, we see again that WMs give an empirical manifestation of BKS:

- the BKS 'contradiction' here is that $\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1 = -1$ (given the PPS of figure 6(*a*)) even though separately $\hat{\sigma}_x^1\hat{\sigma}_y^2 = +1$ and $\hat{\sigma}_x^2\hat{\sigma}_y^1 = +1$;
- these three outcomes can be measured weakly without contradiction because the product of WVs is not equal to the WV of the product;
- if BKS were not correct and a noncontextual HVT were possible, then the product rule should be satisfied and an IM of $\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1$ should yield +1. This leads to an immediate contradiction because
  - by theorem 3, the WV must be equal to the ideal result
  - but, this would be inconsistent with an actual WM which will register $(\hat{\mathbf{N}}_{--})_{\mathrm{w}} = -\frac{1}{2}$,
- therefore, BKS is empirically consistent with WMs.

We have thus given a physical explanation for why an IM cannot reveal these values, while WMs can reveal these values. Thus, with WMs, the BKS 'contradiction' still exists (i.e. a noncontextual HVT cannot reproduce QM), yet now it can also be measured. In other words, we have physically shown how to obtain

- +1 for the product of all nine observables when this is performed in the sequence of the rows of figure 4,
- −1 for the product of all nine observables when this is performed in the sequence of the columns of figure 4

(assuming that the system is PPS and measured weakly). The ambiguity in determining whether $\hat{\sigma}_z^1\hat{\sigma}_z^2 = +1$ is obtained or $\hat{\sigma}_z^1\hat{\sigma}_z^2 = -1$ is obtained gets shifted to the ambiguity of determining which set of boundary conditions is obtained, i.e. it is now a physical property of the system.

As shown in the following section, the WVs calculated in subsection 3.3 (e.g. equations (3.22) and (3.21)) are identical to WVs for EPR entanglement [29] and thus entanglement in a pre-selected state of two particles is isomorphic to an entanglement in our two virtual particles.

3.4. Non-classical weak values for EPR and Peres/BKS

We shall now show that WVs can show new connections between BKS and EPR. Equations (3.22) and (3.21) give the same result as calculating WVs for EPR entanglement [29] and thus there is an interesting new kind of isomorphism between the problems of WVs in BKS nonets and EPR. We can also consider interesting manifestations of the 'contextuality' in these situations by making separate measurements on the ancilla. Consider a pre-selected state which is entangled between one of the particles of the nonet and the ancilla (where, following Mermin, we have chosen the ancilla to be an orthonormal set $|\mathbf{f}^1\rangle$):

$$|A\rangle = \tfrac{1}{\sqrt{2}}\big(\big|\hat{\sigma}_z^1 = +1\big\rangle|\mathbf{f}^0 = +1\rangle - \big|\hat{\sigma}_z^1 = -1\big\rangle|\mathbf{f}^0 = -1\rangle\big). \tag{3.23}$$

If we consider product state post-selections, e.g. with $\hat{\sigma}_x^1 = 1$ and $\mathbf{f}^0 = +1$, then we know that an IM of $\hat{\sigma}_x^1\mathbf{f}^0$ must yield

$$\hat{\sigma}_x^1\mathbf{f}^0 = 1. \tag{3.24}$$

The pre-selected state also yields[6]

$$\hat{\sigma}_x^1\mathbf{f}^0 + \mathbf{f}^1\hat{\sigma}_z^1 = 0. \tag{3.25}$$

Thus

$$\begin{aligned}\left\{\hat{\sigma}_x^1\mathbf{f}^0 + \mathbf{f}^1\sigma_z^1\right\}&\left\{\left|\uparrow_z^1\right\rangle|\mathbf{f}^0 = -1\rangle - \left|\downarrow_z^1\right\rangle|\mathbf{f}^0 = +1\rangle\right\} \\ &= \left\{-\left|\downarrow_z^1\right\rangle|\mathbf{f}^0 = -1\rangle + \left|\uparrow_z^1\right\rangle|\rangle\right\} - \left\{\left|\uparrow_z^1\right\rangle|\rangle - \left|\downarrow_z^1\right\rangle|\mathbf{f}^0 = -1\rangle\right\} = 0.\end{aligned} \tag{3.26}$$

We can deduce outcomes for un-performed measurements of $\hat{\sigma}_z^1$:

$$\hat{\sigma}_z^1 = -1, \tag{3.27}$$

(we have assumed that the ancilla is not disturbed to obtain this conclusion). We can also deduce outcomes for un-performed measurements of $\mathbf{f}^1$ (assuming that the first particle is not disturbed):

$$\mathbf{f}^1 = -1. \tag{3.28}$$

Once again, we see a violation of the product rule [28]: from equation (3.26) we deduce that $\mathbf{f}^1\hat{\sigma}_z^1 = -1$, but if constructed individually from equations (3.28) and (3.27), $\mathbf{f}^1\hat{\sigma}_z^1 = +1$. This 'conclusion', however, relies on counter-factual statements, since not all the required measurements equations (3.25)–(3.28) can be performed simultaneously without disturbing each other. However, we *can* perform WMs on all these statements simultaneously (see subsection 3.3).

WVs in the EPR situation can also be seen in the instant example if we consider this PPS (instead of a Bell state, we measure a definite state of the ancilla). We see the identical non-classical results as seen in the previous section. First, we define the following projectors:

N_{++} the projector on the state $\hat{\sigma}_z^1 = 1$ and $\mathbf{f}^1 = 1$
N_{+-} the projector on the state $\hat{\sigma}_z^1 = 1$ and $\mathbf{f}^1 = -1$
N_{-+} the projector on the state $\hat{\sigma}_z^1 = -1$ and $\mathbf{f}^1 = 1$
N_{--} the projector on the state $\hat{\sigma}_z^1 = -1$ and $\mathbf{f}^1 = -1$.

From theorems 2 and 3, we can deduce the following: from the post-selection $\hat{\sigma}_x^1 = +1$, we can deduce that $\mathbf{f}^1 = -1$, i.e. equation (3.28). The different ways to obtain $\mathbf{f}^1 = -1$ are given by N_{+-} (i.e. $\hat{\sigma}_z^1 = 1$ and $\mathbf{f}^1 = -1$) and N_{--} (i.e. $\hat{\sigma}_z^1 = -1$ and $\mathbf{f}^1 = -1$), and therefore we can deduce the conservation relationship

$$N_{+-} + N_{--} = 1. \tag{3.29}$$

Now $\hat{\sigma}_z^1 = -1$, and thus $\hat{\sigma}_z^1 \neq +1$ is characterized by

$$N_{+*} \equiv N_{++} + N_{+-} = 0. \tag{3.30}$$

In terms of the box analogy, this is how many ways that particle 1 can be found in box 1. In addition, if $\mathbf{f}^1 = -1$ then $\mathbf{f}^1 \neq +1$ thereby giving

$$N_{++} + N_{-+} = 0. \tag{3.31}$$

That is, in how many ways can particle 2 be found in box 2. Furthermore, since equation (3.31) equals equation (3.30), we can deduce

$$N_{+-} = N_{-+}. \tag{3.32}$$

[6] This is easy to see because $\mathbf{f}^1\hat{\sigma}_z^1|\uparrow_z^1\rangle|\mathbf{f}^0 = -1\rangle \to |\uparrow_z^1\rangle|\mathbf{f}^0 = +1\rangle$ and $\hat{\sigma}_x^1\mathbf{f}^0|\uparrow_z^1\rangle|\mathbf{f}^0 = -1\rangle \to -|\downarrow_z^1\rangle|\mathbf{f}^0 = -1\rangle$ and $\mathbf{f}^1\hat{\sigma}_z^1|\downarrow_z^1\rangle|\mathbf{f}^0 = +1\rangle \to -|\downarrow_z^1\rangle|\mathbf{f}^0 = -1\rangle$ and $\hat{\sigma}_x^1\mathbf{f}^0|\downarrow_z^1\rangle|\mathbf{f}^0 = +1\rangle \to |\uparrow_z^1\rangle|\mathbf{f}^0 = +1\rangle$.

From the post-selection, we know that an IM of $\hat{\sigma}_x^1\mathbf{f}^0$ will yield +1, therefore, using equation (3.25), we can deduce that $\mathbf{f}^1\hat{\sigma}_z^1 = -1$ (which means that both particles cannot be found in the same box), i.e.

$$\underbrace{\hat{\sigma}_x^1\mathbf{f}^0}_{=+1} + \underbrace{\mathbf{f}^1\hat{\sigma}_z^1}_{\Rightarrow=-1} = 0. \tag{3.33}$$

Next we ask how to obtain $\mathbf{f}^1\hat{\sigma}_z^1 = -1$. To obtain this, it cannot be that $\mathbf{f}^1\hat{\sigma}_z^1 = +1$ and thus $\mathbf{f}^1$ and $\hat{\sigma}_z^1$ must be opposite in sign. This can be characterized by

$$N_{++} + N_{--} = 0. \tag{3.34}$$

Subtracting equation (3.34) from the following identity:

$$N_{++} + N_{--} + N_{+-} + N_{-+} = +1, \tag{3.35}$$

we obtain

$$N_{+-} + N_{-+} = 1. \tag{3.36}$$

From equations (3.32) and (3.36), we can deduce

$$N_{-+} = N_{+-} = \tfrac{1}{2}. \tag{3.37}$$

Plugging this value into equation (3.30), we can deduce that

$$N_{++} = -\tfrac{1}{2}. \tag{3.38}$$

Finally, plugging this into equation (3.34), we can deduce

$$N_{--} = \tfrac{1}{2}. \tag{3.39}$$

Entanglement in a pre-selected state of two particles is isomorphic to an entanglement in our two virtual particles. Thus, if we look at the right variables, then the BKS setup can be seen to be related to EPR.

The situation analysed above is general and also points to the Peres/BKS-example [46]: in summary, consider a pre-selected state $|\Psi_{\mathrm{EPR}}\rangle = \frac{1}{\sqrt{2}}\left(\left|\uparrow_z^1\right\rangle\left|\downarrow_z^2\right\rangle - \left|\downarrow_z^1\right\rangle\left|\uparrow_z^2\right\rangle\right)$ for which the following identity holds: $\hat{\sigma}_x^1\hat{\sigma}_y^2 + \hat{\sigma}_x^2\hat{\sigma}_y^1 = 0$. In addition, $|\Psi_{\mathrm{EPR}}\rangle$ is also an eigenvector (with eigenvalue -1) of the following operators: $\hat{\sigma}_x^1\hat{\sigma}_x^2, \hat{\sigma}_y^1\hat{\sigma}_y^2, \hat{\sigma}_z^1\hat{\sigma}_z^2$. It is easy to see that a noncontextual HVT cannot assign values consistent with these operator relations

$$V_{\vec{\psi}}\left(\widehat{\sigma_x^1\sigma_x^2}\right) = V_{\vec{\psi}}\left(\widehat{\sigma_y^1\sigma_y^2}\right) = V_{\vec{\psi}}\left(\widehat{\sigma_z^1\sigma_z^2}\right) = -1. \tag{3.40}$$

Consider the commuting observables $\hat{A}_1 = \hat{\sigma}_x^1\hat{\sigma}_y^2$ and $\hat{A}_2 = \hat{\sigma}_x^2\hat{\sigma}_y^1$. We know that

$$\hat{A}_1\hat{A}_2 = \hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1 = \hat{\sigma}_z^2\hat{\sigma}_z^1 = -1 \tag{3.41}$$

in the singlet state. The assumption of noncontextuality is that value assignments can be made to equation (3.41) even when these assignments are taken from a different context: e.g. assigning values from equation (3.40), we obtain

$$V_{\vec{\psi}}\left(\hat{\sigma}_x^1\right)V_{\vec{\psi}}\left(\hat{\sigma}_y^2\right)V_{\vec{\psi}}\left(\hat{\sigma}_x^2\right)V_{\vec{\psi}}\left(\hat{\sigma}_y^1\right) = (-1)(-1) = +1, \tag{3.42}$$

but experimentally, we obtain -1 from equation (3.41), a contradiction. Thus a noncontextual HVT is impossible. However, now we can probe this state by post-selections and obtain $\hat{\sigma}_x^1\hat{\sigma}_y^2 = -1$ even though $\hat{\sigma}_x^1 = -1$ and $\hat{\sigma}_y^2 = -1$. This, again, can only be done with PPSs since in a pre-selected only system, for two commuting observables, the product rule is satisfied in contrast to PPSs. In addition, for this type of contextuality with PPSs, either the pre-selected or the post-selected state must be the EPR state (with an additional assumption of invariance under exchange of particles).

4. Pre-and-post-selection and contextuality in higher dimensions

The feat presented in the previous section can be done in higher dimensions, e.g. to the GHZ state [36, 41]. In this case, we see that it is not possible to replace the spin operators by ordinary numbers (which is what a noncontextual HVT attempts to do). However, assignments can be correctly made to each of the $\hat{\sigma}$'s if post-selection is utilized. Once again, the limitations to these assignments can be seen by using the structure of TSQM: if we try to measure all of these observables together, then some of the values will be assigned in the diagonal-PPS sense, and therefore measuring these observables will cause a disturbance even though they commute. In brief, consider the GHZ case of 3-spins pre-selected in the state

$$|\Psi_{\text{in}}\rangle = \tfrac{1}{\sqrt{2}}\left|\uparrow_z^1\uparrow_z^2\uparrow_z^3\right\rangle - \tfrac{1}{\sqrt{2}}\left|\downarrow_z^1\downarrow_z^2\downarrow_z^3\right\rangle. \tag{4.1}$$

Consider that the pre-selected state is an eigenstate of the following operators: $\hat{A}_1 \equiv \hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_y^3$, $\hat{A}_2 \equiv \hat{\sigma}_y^1\hat{\sigma}_x^2\hat{\sigma}_y^3$ and $\hat{A}_3 \equiv \hat{\sigma}_y^1\hat{\sigma}_y^2\hat{\sigma}_x^3$ with eigenvalue +1, and is an eigenstate of $\hat{A}_4 \equiv \hat{\sigma}_x^1\hat{\sigma}_x^2\hat{\sigma}_x^3$ with eigenvalue -1, and finally $\hat{A}_1\hat{A}_2\hat{A}_3 = -\hat{A}_4$. However, because $\hat{A}_1$, $\hat{A}_2$, $\hat{A}_3$ and $\hat{A}_4$ commute, and because $\hat{\sigma}_x^i$, $\hat{\sigma}_y^j$ and $\hat{\sigma}_y^k$ commute one may ask whether it is possible to also satisfy the above equations by replacing the spin operators by ordinary numbers $\hat{\sigma}_x^1 = \pm 1$, $\hat{\sigma}_y^1 = \pm 1$. Without post-selection, this cannot be done because the assignments are, once again, inconsistent with the multiplicative structure of the observables because $\hat{A}_1\hat{A}_2\hat{A}_3 = 1$, a contradiction.

4.1. GHZ and PPS

Assignments can be correctly made to each of the $\hat{\sigma}$'s if post-selection is utilized with limitations again arising from diagonal assignments. For example, consider a post-selection of the three particles by measuring their x component with $\hat{\sigma}_x = -1$:

$$\langle\Psi_{\text{fin}}| = \left\langle\downarrow_x^1\downarrow_x^2\downarrow_x^3\right|. \tag{4.2}$$

For example, from the post-selected state we *strongly* know that $\hat{\sigma}_x^1\hat{\sigma}_x^2\hat{\sigma}_x^3 = -1$ and from the pre-selected state we *strongly* know that $\hat{\sigma}_x^i\hat{\sigma}_y^j\hat{\sigma}_y^k = +1$. Thus statements of the form $\hat{\sigma}_y^j\hat{\sigma}_y^k = -1$ can only be made when information is used from both the pre-selected vector *and* from the post-selected vector. For example, if a measurement of $\hat{\sigma}_y^1\hat{\sigma}_y^2$ is performed, then we will definitely find $\hat{\sigma}_y^1\hat{\sigma}_y^2 = -1$. However, if we attempt to perform a second measurement, e.g. of $\hat{\sigma}_y^3\hat{\sigma}_y^1$ then we will not find $\hat{\sigma}_y^3\hat{\sigma}_y^1 = -1$, because the $\hat{\sigma}_y^1\hat{\sigma}_y^2$ measurement will destroy the pre-selected vector which contains the $\hat{\sigma}_x^i\hat{\sigma}_y^j\hat{\sigma}_y^k = +1$ information that the $\hat{\sigma}_y^3\hat{\sigma}_y^1$ measurement depends on. This disturbance occurs even though $\hat{\sigma}_y^1\hat{\sigma}_y^2$ and $\hat{\sigma}_y^3\hat{\sigma}_y^1$ commute!

4.2. Weak values in GHZ state

We may also consider WMs of the GHZ observables [67]. With the post-selection $\langle\Psi_{\text{fin}}| = \left\langle\downarrow_x^1\downarrow_x^2\downarrow_x^3\right|$, then in the intermediate time we can replace $\hat{\sigma}_x^1 = \hat{\sigma}_x^2 = \hat{\sigma}_x^3 = -1$ and taking the inner product with this post-selection $|\Psi_{\text{fin}}\rangle$, we then find

$$\left(\hat{\sigma}_y^2\hat{\sigma}_y^3\right)_{\text{w}} \equiv \frac{\langle\Psi_{\text{fin}}|\hat{\sigma}_y^2\hat{\sigma}_y^3|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{in}}\rangle} = -1 \tag{4.3}$$

$$\left(\hat{\sigma}_y^1\hat{\sigma}_y^2\right)_{\text{w}} \equiv \frac{\langle\Psi_{\text{fin}}|\hat{\sigma}_y^1\hat{\sigma}_y^2|\Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|\Psi_{\text{in}}\rangle} = -1 \tag{4.4}$$

$$\left(\hat{\sigma}_y^1\hat{\sigma}_y^3\right)_{\rm w} \equiv \frac{\langle\Psi_{\rm fin}|\hat{\sigma}_y^1\hat{\sigma}_y^3|\Psi_{\rm in}\rangle}{\langle\Psi_{\rm fin}|\Psi_{\rm in}\rangle} = -1. \tag{4.5}$$

Using again the analogy with particles in boxes, equation (4.3) means that particles 2 and 3 are not together in the same box, while equation (4.4) means that particles 1 and 2 are not together in the same box, and equation (4.5) means that particles 1 and 3 are not together in the same box. But we only have two boxes, so if 1 and 2 are not in the same box and 1 and 3 are not in the same box, then 2 and 3 must be in the same box. It is clear that the above equalities cannot be satisfied simultaneously by replacing the operators for classical numbers taking the values ± 1. To simplify this analysis, we define

$$(\hat{\mathbf{N}}_{+++})_w = |\uparrow_y\rangle\langle\uparrow_y|_1 \otimes |\uparrow_y\rangle\langle\uparrow_y|_2 \otimes |\uparrow_y\rangle\langle\uparrow_y|_3, \tag{4.6}$$

where the two boxes are denoted by $\pm$ referring to the spin component along y. Using theorems 2 and 3, it can be shown that

$$(\hat{\mathbf{N}}_{+++})_w = \hat{\mathbf{N}}_{---} = -\tfrac{1}{4}, \tag{4.7}$$

and

$$(\hat{\mathbf{N}}_{++-})_w = (\hat{\mathbf{N}}_{--+})_w = (\hat{\mathbf{N}}_{+-+})_w = (\hat{\mathbf{N}}_{-+-})_w = \cdots = \tfrac{1}{4}. \tag{4.8}$$

We have thus shown how to obtain non-classical WVs, i.e. negative triplet occupations.

5. Discussion

In this paper, we first emphasized the purely theoretical aspect of WVs by themselves before considering their realization in actual measurement situations (i.e. in WMs). Although WVs require an ensemble in order to be probed empirically with WMs, we can consider WVs by themselves as a general, non-statistical and robust mathematical property of every individual PPS system (see appendix B). By first focusing on the logical or mathematical aspects of WVs, we were able to obtain novel insights on the subject of contextuality. For example:

- We proved that if we start with a situation that is 'primed to reflect contextuality' as reflected in BKS, then we can always find a post-selection which can empirically manifest non-classical WVs. That is, with every pre- and post-selection leading to 'quantum contextuality,' some WVs will have a unique signature, namely a WV outside of the eigenvalue spectrum. We believe that it is precisely these WVs which cannot be reproduced by a noncontextual HVT (see subsections 5.1 and 5.2 below). Analysis of the WVs shows how QM itself (rather than HVTs) can *cope* with the apparent paradox of contextual situations in a new and interesting way. As Mermin emphasized: '. . . what follows is not idle theorizing about 'hidden variables'. It is a rock solid quantum mechanical effort to answer a perfectly legitimate quantum mechanical question' [26].
- The breakdown of noncontextuality in the three-box and Mermin examples required the system to be pre- and post-selected from the beginning. In the three-box paradox of subsection 1.3:
  - $\hat{\mathbf{P}}_A = 1$ if only box A is opened, while $\hat{\mathbf{P}}_B = 1$ if only box B is opened,
  - but if we measure both boxes A and B, then the particle will not be found in both boxes, i.e. $\hat{\mathbf{P}}_A\hat{\mathbf{P}}_B = 0$ even though $\hat{\mathbf{P}}_A$ and $\hat{\mathbf{P}}_B$ commute, a violation of the product rule.
  - If WMs are performed, then the non-classical result $(\hat{\mathbf{P}}_C)_{\rm w} = -1$ is obtained.

- With respect to the Mermin example:
  - 'quantum contextuality' is demonstrated here by $\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1 = -1$ (given the PPS of figure 4(*a*)) even though separately $\hat{\sigma}_x^1\hat{\sigma}_y^2 = +1$ and $\hat{\sigma}_x^2\hat{\sigma}_y^1 = +1$;
  - these three values can co-exist because the product of WVs is not equal to the WV of the product;
  - if 'quantum contextuality' were not correct and a noncontextual assignment of values to operators were possible, then the product rule should be satisfied and $\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1$ should yield +1. This leads to an immediate contradiction because
    * by theorem 2, the WV must be equal to the ideal result,
    * but this would be inconsistent because the calculated WV is $\left(\hat{\sigma}_x^1\hat{\sigma}_y^2\hat{\sigma}_x^2\hat{\sigma}_y^1\right)_{\mathrm{w}} = -1$.
- However, the breakdown of noncontextuality in GHZ and EPR can be seen with just a pre-selection by itself, i.e. noncontextual HV assignments cannot be made. We are therefore adding a new, previously un-explored perspective by showing that post-selections suggest a physical picture for why the assignments cannot be made. In addition, the existence of strange WVs demonstrates a new way that QM copes with contextuality. For example, in the 4D/EPR case, $\hat{\sigma}_x^1\hat{\sigma}_y^2 = -1$ even though $\hat{\sigma}_x^1 = -1$ and $\hat{\sigma}_y^2 = -1$. Before this development, the 'contextuality' in these examples was thought only to exist at the level of HVTs, but now it can be probed empirically.
- We have also demonstrated a new approach to contextuality in PPS situations with just IMs because the sum rule is not necessarily satisfied. The reason follows the same interference argument used for the violation of the product rule with WVs.
- Finally, as we argued in [11], WVs obey a simple, intuitive and, most importantly, *self-consistent* logic. This is in stark contrast with the logic of the original counter-factual statements which is not internally self-consistent and leads into paradoxes. We are convinced that, due to its self-consistency, the WV logic will lead to a deeper understanding of the nature of quantum mechanics. To illustrate this intuition again, by way of example, consider a generic multi-particle system in a higher dimensional space $\rho(x_1, x_2) \sim \Psi_{\mathrm{fin}}^*(x_1, x_2)\Psi_{\mathrm{in}}(x_1, x_2)$ and also consider projections onto space and time. In general, $\rho(x_1, x_2)$ cannot be measured locally because there is no local way of measuring particle 1 at x_1 simultaneously with measuring particle 2 at x_2: this would involve a nonlocal Hamiltonian. However, the projection on each line can be measured separately: if it is known that particle 2 is at x_2 then this means $\int \rho(x_1, x_2)\,\mathrm{d}x_1 = 1$. Similarly, if it is known that particle 1 is at x_1 then this means $\int \rho(x_1, x_2)\,\mathrm{d}x_2 = 1$. However, asking the question if *both* particles are there together is another point in phase space (see figure 11). Our main point is that the particle also could have been located simultaneously in another position with certainty, but then we would have to place a negative number somewhere else in order to satisfy the global constraint, i.e. the integral of the WVs has to add up to 1 (because there is just a single particle along each line), but the individual numbers at each point can be arbitrary. IM outcomes reveal an integration along just one line, but the WV densities are not just those lines. For example, in the Hardy case [11] the seeming contradiction that both particles are there individually but are not there together is resolved by a negative number of particles at another point in phase space.
- Finally, WVs in logical PPS-paradoxes can be revealed (under certain circumstances) through WMs (see appendix B). We have therefore shown new ways that the 'charming elementary mathematics' [26] of BKS can manifest empirically.

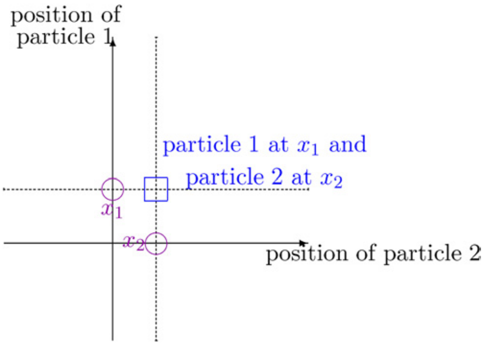

Figure 11. Weak value density for two particles.

5.1. Possible relationships between HVTs and weak values

We summarize some general considerations on possible relationships between HVTs and WVs.

5.1.1. New criterion for HVTs to reproduce weak values. Let us consider possible HVs for the particular WM and WV described in appendix A. Suppose for a single particle, we have a HV which gives the pre-selection, $|\uparrow_x\rangle$, another HV which gives the post-selection, $|\uparrow_y\rangle$, and another HV which gives the result for an intermediate IM of $\frac{\hat{\sigma}_x+\hat{\sigma}_y}{\sqrt{2}}$ which could be +1 or -1. But as shown in appendix B, a WM for an ensemble of such particles will robustly reveal a WV of $\sqrt{2}$. How then could we associate a HV for this WM? There are two general possibilities.

- If the WV is a property of an individual system (not just an ensemble, see appendix B) and if in addition we want to add HVs, then it appears that the 2 are inconsistent unless any WM during the intermediate time will definitely disturb the HV (because the IM yields a different result).
- A WM must give a different value than IMs because a WM cannot be performed robustly on a single system.

The first option shows that the HVs which are supposed to give the outcome for all IMs become completely sensitive to an arbitrarily weak WM. Indeed, even in the limit of arbitrarily weak interactions, there will still be a finite disturbance to the HV for the same large fraction of particles in the ensemble. For a number of reasons (see, e.g. appendix B), the second option is also, in our opinion, not very physical. Nevertheless, these considerations suggest new criterion for HVs in addition to contextuality and non-locality. That is one or both of the following assumptions would have to be violated.

- *Criterion 1.* HVs should be stable, i.e. an arbitrarily small disturbance should not disturb the HVs.
- *Criterion 2.* If a HV has a certain value, then if anything couples weakly to it, then the weak coupling should produce a completely different value.

Finally, it was also argued [24, 29] that a noncontextual HVT can reproduce QM if we allow for a disturbance of the HVs: 'the possibility of measurement disturbance blocks the conclusion that a PPS-paradox is itself a proof of the contextuality of HVTs'. The motivation

[24] behind this assertion was the belief that PPS-paradoxes could be explained entirely within classical mechanics [39, 40] and that contextuality should not be regarded as fundamental in a classical picture of reality. Nevertheless, the ‘paradoxical’, i.e. non-classical, nature of the three-box paradox was recently reaffirmed in terms of IMs [42]. In this paper, we expanded this point in a number of ways, e.g. by arguing that non-classical WVs can be empirically demonstrated in PPS-paradoxes. Non-classical WVs include those outside the eigenvalue spectrum which cannot be reproduced from any positive definite probability distribution of eigenvalues [1, 29, 30]. In addition, with WMs, there is no measurement disturbance, yet non-classical results are still obtained. If we consider robust WMs in which there is no disturbance to the quantum state (see appendix B and [54]), then if we assume criterion 1, i.e. that a WM does not considerably modify the HV, then a HV proof of contextuality (rather than quantum contextuality) could proceed. This would therefore weaken the conclusion that ‘the possibility of measurement disturbance blocks the conclusion that a PPS-paradox is itself a proof of the contextuality of HVTs’ [25] while strengthening the proof of contextuality [24].

5.1.2. Disharmony between quantum measurement theory and HVTs. Even before any consideration of WMs, WVs, or PPS, we suggest another general difficulty for HVTs. Many HVT approaches [24] require that non-commuting observables (such as p and x) can have a ‘simultaneous’ precise reality, as suggested, e.g. by the Wigner–Moyal method. If we require that any theoretical formalism should include exactly what can be measured (no more and no less), then it should be possible to make measurements on these projections. While such densities do give the correct average of a function, i.e. $\int \rho(x,p)f(x,p)\,\mathrm{d}x\,\mathrm{d}p$ (thus appear to behave as proper densities), they also have un-physical aspects, i.e. mathematical artefacts, when the densities become negative. The reason (as will be shown subsequently) is that if we attempt to actually measure such ‘negative’ properties, then the result does not correspond to a physical observable in Hilbert space. For example, if we did try to project on p and x as densities simultaneously, then we obtain the parity operator, taking a generic $\psi(x)$ to $\psi(-x)$. To see this, we translate the classical projection $p = 0$ and $x = 0$ into QM:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\alpha x+\mathrm{i}\beta p}\,\mathrm{d}\alpha\,\mathrm{d}\beta \underbrace{\Rightarrow}_{QM} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{\frac{\mathrm{i}\alpha\beta}{2}}\,\mathrm{e}^{\mathrm{i}\alpha\hat{x}}\,\mathrm{e}^{\mathrm{i}\beta\hat{p}}\,\mathrm{d}\alpha\,\mathrm{d}\beta. \tag{5.1}$$

Consider applying this to a generic wavefunction. First, the exponential, $\mathrm{e}^{\mathrm{i}\beta\hat{p}}$, translates $\psi(x)$. Integrating then over α produces a delta function

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{i\frac{\alpha\beta}{2}}\,\mathrm{e}^{\mathrm{i}\alpha x}\,\psi(x+\beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta = \int_{-\infty}^{\infty}\left\{\underbrace{\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}\alpha(x+\frac{\beta}{2})}\,\mathrm{d}\alpha}_{\delta(x+\frac{\beta}{2})}\right\}\psi(x+\beta)\,\mathrm{d}\beta. \tag{5.2}$$

Finally, integrating over β, we obtain $\beta = -2x$, and thus $\psi(x-2x) = \psi(-x)$. Therefore, the quantum analogue of the classical projection does not correspond to a quantum projector: it corresponds to a highly non-local result, the parity operator.

Therefore, while there is an operational meaning to a density over a set of commuting observables, there are significant difficulties with densities over a set of non-commuting observables even though such densities may have formal utility as an aide to calculation.

5.2. Open question: can noncontextual HVTs account for weak values?

It remains a significant open question whether any noncontextual model, such as those proposed in [24], can reproduce the WV results discussed in this paper. While there is no proof, the uniqueness of quantum-WVs and quantum-WMs in the three-box paradox has been addressed separately in [42] and experimentally in [34]. It has been demonstrated in [1] that strange WVs cannot be reproduced from any positive definite probability distribution of eigenvalues. For this and other reasons (to be addressed in a future paper), we therefore believe that strange WVs cannot be reproduced in a classical-like HVT without pre- and post-selection and without quantum interference.

6. Conclusion

We have analysed contextuality in terms of pre- and post-selection, and have shown that it is possible to assign definite values to observables in new and surprising ways. We presented new physical reasons for restrictions on these assignments. WMs suggest that novel *experimental* aspects of contextuality can be empirically demonstrated. We also proved that every logical PPS-paradox directly implies 'quantum contextuality' which is introduced as the analogue of contextuality at the level of QM rather than at the level of HVTs. Finally, we argued that certain results of these measurements (e.g. eccentric weak values outside the eigenvalue spectrum) cannot be explained by a 'classical-like' HVT.

Although the outcomes of the WMs suggest a story which appears to be even stranger than the original one, the situation is in fact far better. The WVs obey a simple, intuitive and, most importantly, *self-consistent* logic. This is in stark contrast with the logic of the original counter-factual statements which is not internally self-consistent and leads into paradoxes. Strangeness by itself is not a problem; self-consistency is the real issue. In this sense the logic of the WVs is similar to the logic of special relativity. That light has the same velocity in all reference frames is certainly highly unusual, but everything works in a self consistent way, and because of this special relativity is rather easy to understand. We are convinced that, due to its self-consistency, the WMs logic will lead to a deeper understanding of the nature of QM.

Acknowledgments

The author thanks Yakir Aharonov and Sandu Popescu for many conversations and NIST for financial support. The author also thanks J Gray and AD Parks for editing help.

Appendix A. Information gain without disturbance: safety in numbers

It follows from equation (1.9) that the probability for a collapse decreases as $O(\lambda^2)$, but the measuring device's shift grows linearly $O(\lambda)$, so $\delta P_{\mathrm{md}} = \lambda a_i$. For a sufficiently weak interaction (e.g. $\lambda \ll 1$), the probability for a collapse can be made arbitrarily small, while the measurement still yields information but becomes less precise because the shift in the measuring device is much smaller than its uncertainty $\delta P_{\mathrm{md}} \ll \Delta P_{\mathrm{md}}$ (figure 3(*b*)). If we perform this measurement on a single particle, then two non-orthogonal states will be indistinguishable. If this were possible, it would violate unitarity because these states could time evolve into orthogonal states $|\Psi_1\rangle\big|\Phi^{\mathrm{in}}_{\mathrm{md}}\big\rangle \longrightarrow |\Psi_1\rangle\big|\Phi^{\mathrm{in}}_{\mathrm{md}}(1)\big\rangle$ and $|\Psi_2\rangle\big|\Phi^{\mathrm{in}}_{\mathrm{md}}\big\rangle \longrightarrow |\Psi_2\rangle\big|\Phi^{\mathrm{in}}_{\mathrm{md}}(2)\big\rangle$, with $|\Psi_1\rangle\big|\Phi^{\mathrm{in}}_{\mathrm{md}}(1)\big\rangle$ orthogonal to $|\Psi_2\rangle\big|\Phi^{\mathrm{in}}_{\mathrm{md}}(2)\big\rangle$. With weakened measurement interactions, this does not happen because the measurement of these two non-orthogonal states causes a smaller shift in the measuring device than its uncertainty. We conclude that the shift δP_{md} of the measuring device is a measurement error because $\tilde{\Phi}^{\mathrm{MD}}_{\mathrm{fin}}(P_{\mathrm{md}}) = \big\langle P_{\mathrm{md}} - \lambda\langle\hat{A}\rangle\big|\Phi^{\mathrm{in}}_{\mathrm{md}}\big\rangle \approx \big\langle P_{\mathrm{md}}\big|\Phi^{\mathrm{in}}_{\mathrm{md}}\big\rangle$ for $\lambda \ll 1$. Nevertheless, if a large $\big(N \geqslant \frac{N'}{\lambda}\big)$ ensemble of particles is used, then the shift of all the measuring devices $\big(\delta P^{\mathrm{tot}}_{\mathrm{md}} \approx \lambda\langle\hat{A}\rangle\frac{N'}{\lambda} = N'\langle\hat{A}\rangle\big)$

becomes distinguishable because of repeated integrations, while the collapse probability still goes to zero. That is, for a large ensemble of particles which are all either $|\Psi_2\rangle$ or $|\Psi_1\rangle$, this measurement can distinguish between them even if $|\Psi_2\rangle$ and $|\Psi_1\rangle$ are not orthogonal (because the scalar product $\left\langle\Psi_1^{(N)}\middle|\Psi_2^{(N)}\right\rangle = \cos^n\theta \longrightarrow 0$).

Using these observations, we now emphasize that the average of any operator $\hat{A}$, i.e. $\langle\hat{A}\rangle \equiv \langle\Psi|\hat{A}|\Psi\rangle$, can be obtained in three distinct cases [30].

(i) *Statistical method with disturbance*. The traditional approach is to perform IMs of $\hat{A}$ on each particle, obtaining a variety of different eigenvalues, and then manually calculate the usual statistical average to obtain $\langle\hat{A}\rangle$.

(ii) *Statistical method without disturbance*. As demonstrated by using $\hat{A}|\Psi\rangle = \langle\hat{A}\rangle|\Psi\rangle + \Delta A|\Psi_\perp\rangle$. We can also verify that there was no disturbance: consider the spin-1/2 arrangement in the Mermin example (subsubsection 3.2.1), pre-selecting an ensemble, $|\uparrow_x\rangle$, then performing a weakened measurement of $\hat{\sigma}_\xi$ and finally a post-selection again in the x-direction (figure B1). For every post-selection, we will again find $|\uparrow_x\rangle$ with greater and greater certainty (in the weakness limit), verifying our claim of no disturbance. Each measuring device is centred on $\langle\uparrow_x|\sigma_\xi|\uparrow_x\rangle = \frac{1}{\sqrt{2}}$ and the whole ensemble can be used to reduce the spread. The weakened interaction for $\hat{\sigma}_\xi$ means that the inhomogeneity in the magnetic field induces a shift in momentum which is less than the uncertainty $\delta P_{\mathrm{md}}^{\xi} < \Delta P_{\mathrm{md}}^{\xi}$, and thus a wave packet corresponding to $\frac{\hat{\sigma}_x+\hat{\sigma}_y}{\sqrt{2}} = 1$ will be broadly overlapping with the wave packet corresponding to $\frac{\hat{\sigma}_x+\hat{\sigma}_y}{\sqrt{2}} = -1$.

(iii) *Non-statistical method without disturbance*. It is the case where $\langle\Psi|\hat{A}|\Psi\rangle$ is the 'eigenvalue' of a single 'collective operator,' $\hat{A}^{(N)} \equiv \frac{1}{N}\sum_{i=1}^{N}\hat{A}_i$ (with $\hat{A}_i$ being the same operator $\hat{A}$ acting on the ith particle). Using this, we are able to obtain information about $\langle\Psi|\hat{A}|\Psi\rangle$ without causing disturbance (or a collapse) and without using a statistical approach because any product state $|\Psi^{(N)}\rangle$ becomes an eigenstate of the operator $\hat{A}^{(N)}$. To see this, we apply theorem 1 ($\hat{A}|\Psi\rangle = \langle\hat{A}\rangle|\Psi\rangle + \Delta A|\Psi_\perp\rangle$ see footnote 4) to $\hat{A}^{(N)}|\Psi^{(N)}\rangle$, i.e.,

$$\hat{A}^{(N)}|\Psi^{(N)}\rangle = \frac{1}{N}\left[N\langle\hat{A}\rangle|\Psi^{(N)}\rangle + \Delta A\sum_i\left|\Psi_\perp^{(N)}(\mathrm{i})\right\rangle\right], \tag{A.1}$$

where $\langle\hat{A}\rangle$ is the average for any one particle and the states $\left|\Psi_\perp^{(N)}(\mathrm{i})\right\rangle$ are mutually orthogonal and are given by $\left|\Psi_\perp^{(N)}(\mathrm{i})\right\rangle = |\Psi\rangle_1|\Psi\rangle_2\cdots|\Psi_\perp\rangle_i\cdots|\Psi\rangle_N$. That is, the ith state has particle i changed to an orthogonal state and all the other particles remain in the same state. If we further define a normalized state $\left|\Psi_\perp^{(N)}\right\rangle = \sum_i\frac{1}{\sqrt{N}}\left|\Psi_\perp^{(N)}(\mathrm{i})\right\rangle$ then the last term of equation (A.1) is $\frac{\Delta A}{\sqrt{N}}\left|\Psi_\perp^{(N)}\right\rangle$ and its size is $\left|\frac{\Delta A}{\sqrt{N}}\left|\Psi_\perp^{(N)}\right\rangle\right|^2 \propto \frac{1}{N} \rightarrow 0$. Therefore, $|\Psi^{(N)}\rangle$ becomes an eigenstate of $\hat{A}^{(N)}$, with the value $\langle\hat{A}\rangle$ and not even a single particle has been disturbed (as $\hat{N}\rightarrow\infty$).

Tradition has dictated that when measurement interactions are limited so there is no disturbance on the system, then no information can be gained. However, we have now shown that when considered as a limiting process, the disturbance goes to zero more quickly than the shift in the measuring device, which means for a large enough ensemble, information (e.g. the expectation value) can be obtained even though not even a single particle is disturbed.

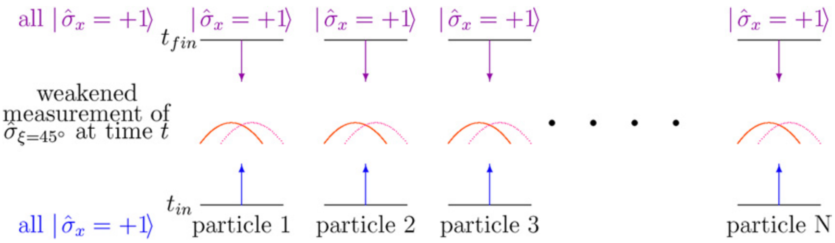

Figure B1. Obtaining the average for an ensemble.

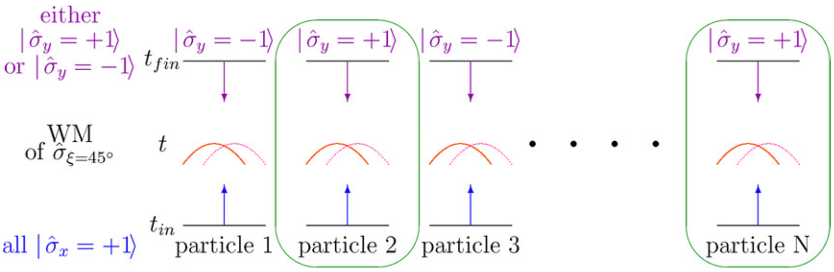

Figure B2. Obtaining the WV statistically for an ensemble.

Appendix B. Weak values can be considered as logical properties of individual systems

For the last case, we can also verify that there was no disturbance caused to the system. By way of example, suppose we were to apply the last option to aspects of the Mermin example (subsection 3.3). Suppose we were to pre-select each particle in $x, |\uparrow_x\rangle$, (i.e. $\left|\Psi_{\text{in}}^{(N)}\right\rangle = \prod_{\text{j}=1}^{N} |\uparrow_x\rangle_j$), perform a measurement of the collective observable in the 45° angle to the x–y plane of $\hat{\sigma}_\xi^{(N)} \equiv \frac{1}{N}\sum_{\text{i}=1}^{N}\hat{\sigma}_\xi^i$, and then perform a post-selection again in the x-direction. For every post-selection, we will again find $|\uparrow_x\rangle$ with certainty, verifying our claim of no disturbance. Over the whole ensemble, the measuring device will robustly register $\langle\hat{\sigma}_\xi\rangle = \frac{1}{\sqrt{2}}$, i.e. the average for an individual particle, which is the same answer as the first (statistical) method. This is thus consistent with the statement that the *average is a property of the individual particle*, only its value is obtained robustly (by summing the collective operator) in a non-statistical sense over the whole ensemble (see figure B1).

To see how this is relevant for WMs and WVs, once again, we insert a complete set of states $\{|\Psi_{\text{fin}}\rangle_j\}$ into $\langle\Psi_{\text{in}}|\hat{A}|\Psi_{\text{in}}\rangle$ (where we take $|\Psi_{\text{in}}\rangle \equiv |\uparrow_x\rangle$) yielding equation (1.10). We interpret the states $\langle\Psi_{\text{fin}}|_j$ as the possible outcomes of a final ideal measurement on the system, i.e. a post-selection. This is precisely how we interpreted $\langle\Psi_{\text{in}}|$ in the verification of the expectation value. We can therefore analyse the same problem with a post-selection different from the original definite $\langle\Psi_{\text{in}}|$ post-selection used in verifying the expectation value. By way of example, let us consider a post-selection in the y-direction, i.e. we have N particles pre-selected with $\hat{\sigma}_x = 1$, a WM of $\hat{\sigma}_{45}$ (which does not significantly disturb the spins which are thus still in the state $\hat{\sigma}_x = 1$ after the $\hat{\sigma}_{45}$ measurement) and followed by a post-selection in the y-direction. The probability of obtaining $\hat{\sigma}_y = 1$ is $1/2$ and thus the total probability of finding all N spins with $\hat{\sigma}_y = 1$ is an exponentially small 2^{-N} (see figure B2).

If we look at only those results with $\hat{\sigma}_y = 1$, then the pointer is robustly shifted by the strange WV $\sqrt{2}$:

$$(\hat{\sigma}_\xi)_{\mathrm{w}} = \frac{\prod_{k=1}^{N} \langle \uparrow_y |_k \sum_{\mathrm{i}=1}^{N} \left\{ \hat{\sigma}_x^i + \hat{\sigma}_y^i \right\} \prod_{\mathrm{j}=1}^{N} | \uparrow_x \rangle_j}{\sqrt{2} N (\langle \uparrow_y | \uparrow_x \rangle)^N} = \sqrt{2} \pm O\left(\frac{1}{\sqrt{N}}\right). \tag{B.1}$$

Using the same argument as for the expectation value, WVs can be considered as a non-statistical robust mathematical property of an ensemble. However, if we do not see the strange WV, this does not mean that the WV was not already 'there,' only we were not lucky enough to see it. That is, in the first case of verifying the expectation value, we knew that every member of the ensemble would satisfy the post-selection of $\langle \uparrow_x |$. With different post-selections, the only difference is that we do not know ahead of time which sub-ensemble will satisfy a particular post-selection criteria. But this does not mean that the WV cannot be analysed separately from actual experiments. One might argue that the third case, i.e. the non-statistical method without disturbance, is not sufficiently general because the probability $|\langle \Psi_{\mathrm{fin}} | \Psi_{\mathrm{in}} \rangle|^{2N}$ for all N particles to end up in the same final state $|\Psi_{\mathrm{fin}}\rangle$ becomes exponentially small. Recently [30] it was shown how to obtain a robust WV with a much smaller sample: e.g. for the particular pre- and post-selection considered here, approximately $N/2$ out of N pre-selected particles will satisfy the post-selection criterion and thus this result is not a rare outcome.

References

[1] Aharonov Y and Bohm D 1961 *Phys. Rev.* **122** 1649
Wheeler J A and Zurek W H (eds) 1983 *Quantum Theory and Measurement* (Princeton, NJ: Princeton University Press) (reprint)
Aharonov Y, Pendelton H and Petersen A 1969 *Int. J. Theor. Phys.* **2** 213
[2] Aharonov Y, Bergmann P G and Lebowitz J L 1964 *Phys. Rev.* **134** B1410
Wheeler J A and Zurek W H (eds) 1983 *Quantum Theory and Measurement* (Princeton, NJ: Princeton University Press) pp 680–6 (reprint).
[3] Albert D, Aharonov Y and D'Amato S 1985 *Phys. Rev. Lett.* **54** 5
[4] Vaidman L, Aharonov Y and Albert D 1987 *Phys. Rev. Lett.* **58** 1385
[5] Aharonov Y and Vaidman L 1990 *Phys. Rev.* A **41** 11
[6] Aharonov Y and Vaidman L 1991 *J. Phys. A: Math. Gen.* **24** 2315
[7] Aharonov Y and Vaidman L 1993 *Phys. Lett.* A **178** 38
[8] Aharonov Y, Davidovich L and Zagury N 1993 *Phys. Rev.* A **48** 1687
[9] Reznik B and Aharonov Y 1995 *Phys. Rev.* A **52** 2538
[10] Aharonov Y, Massar S, Popescu S, Tollaksen J and Vaidman L 1996 *Phys. Rev. Lett.* **77** 983
[11] Aharonov Y, Botero A, Popescu S, Reznik B and Tollaksen J 2002 *Phys. Lett.* A **301** 130–8
[12] Aharonov Y and Englert B-G 2001 *Z. Naturforsch.* A **56** 16–9
[13] Aharonov Y and Rohrlich D 2005 *Quantum Paradoxes: Quantum Theory for the Perplexed* (Cambridge: Wiley)
[14] Aharonov Y and Botero A 2005 *Phys. Rev.* A **72** 052111
[15] Botero A and Reznik B 2000 *Phys. Rev.* A **61** 050301-4
[16] Bub J 2001 *Phys. Rev.* A **63** 032309
[17] Englert B-G, Kurtsiefer C and Weinfurter H *Phys. Rev.* A **63** 032303
[18] Beige A, Englert B-G, Kurtsiefer C and Weinfurter H 2001 *Preprint* quant-ph/0101066
[19] Bell J S 1964 *Physics* **1** 195–200
[20] Bell J S 1966 *Rev. Mod. Phys.* **38** 447–52
Bell J S 1987 *Speakable and Unspeakable in Quantum mechanics* (Cambridge: Cambridge University Press) (reprint)
[21] Chiao R Y and Steinberg A M 1997 *Prog. Opt.* **37** 345–405
Solli D R *et al* 2004 *Phys. Rev. Lett.* **92** 043601
[22] Gell-Mann M and Hartle J B 1993 *Phys. Rev.* D **47** 3345
[23] Kochen , Simon , Specker and Ernst 1967 *J. Math. Mech.* **17** 59–87
[24] Leifer M S and Spekkens R W 2005 *Int. J. Theor. Phys.* **44** 1977–87 (*Preprint* quant-ph/0412179)
[25] Leifer M S and Spekkens R W 2005 *Phys. Rev. Lett.* **95** 200405

[26] Mermin N D 1995 *Phys. Rev. Lett.* **74** 831
Cohen R S (ed) *et al* 1997 *Potentiality, Entanglement and Passion-at-a-Distance* (Dordrecht: Kluwer) pp 149–57
[27] Mermin N D 1993 *Rev. Mod. Phys.* **65** 803
[28] Vaidman L 1993 *Phys. Rev. Lett.* **70** 3369
Lahti P J, Bush P and Mittelstaedt P (eds) 1993 *Symp. Found. Mod. Phys.* 3 406–17
[29] Tollaksen J 2001 *PhD Thesis* Boston University, Y Aharonov (thesis advisor)
[30] Tollaksen J 2007 Robust weak measurements on finite samples *Journal of Physics Conference Series vol 70* 012015 (*Preprint* quant-ph/0703038)
[31] Aharonov Y and Gruss E 2005 *Preprint* quant-ph/0507269
[32] Gleason A 1957 *J. Math. Mech.* **6** 885
[33] Steinberg A M 1995 *Phys. Rev. Lett.* **74** 2405
[34] Resch K J, Lundeen J S and Steinberg A M 2004 *Phys. Lett.* A **324** 125–31
[35] Lundeen J S, Resch K J and Steinberg A M 2005 *Phys. Rev.* A **72** 016101 part A–B
Ahnert S E and Payne M C 2004 *Phys. Rev.* A **70** 042102
[36] Vaidman L 1996 *Found. Phys.* **26** 895
[37] von Neumann J 1983 *Mathematical Foundations of Quantum Theory* (Princeton, NJ: Princeton University Press)
[38] Simon C, Zukowski M, Weinfurter H and Zeilinger A 2000 *Phys. Rev. Lett.* **85** 1783–6
[39] Kastner R *Found. Phys.* **29** 851–63
[40] Kirkpatrick K 2003 *J. Phys. A: Math. Gen.* **36** 4891–900
[41] Greenberger D M, Horne M A and Zeilinger A 1989 *Bell's Theorem, Quantum Theory and Conceptions of the Universe* ed M Kafatos (Dordrecht: Kluwer) p 69
See also Greenberger D M, Horne M A, Shimony A and Zeilinger A 1990 *Am. J. Phys.* **58** 1131–43
[42] Ravon T and Vaidman L 2006 The 3-box paradox re-visited *Preprint* quant-ph/0606067
[43] Hardy L 1992 *Phys. Rev. Lett.* **68** 2981
[44] Bub J and Brown H 1986 *Phys. Rev. Lett.* **56** 2337
[45] Oreshkov O and Brun T A 2005 *Phys. Rev. Lett.* **95** 110409
[46] Peres A 1991 *J. Phys. A: Math. Gen.* **24** L175
[47] Feynman R Negative probability *Quantum Implications* ed B J Hiley and F D Peat (New York: Routledge and Kegan Paul) pp 224–6
[48] Spekkens R W 2005 Contextuality for preparations, transformations, and unsharp measurements *Phys. Rev.* A **71** 052108
[49] Aharonov Y and Vaidman L 2001 *Time in Quantum Mechanics* ed J Muga, R S Mayato and I Egusquiza (*Preprint* quant-ph/0105101)
[50] Aharonov Y and Tollaksen J 2007 New insights on time-symmetry in quantum mechanics *Visions of Discovery: New Light on Physics, Cosmology, and Consciousness* ed R Y Chiao, M L Cohen, A J Leggett, W D Phillips and C L Harper Jr (Cambridge: Cambridge University Press)
[51] Tollaksen J 2007 Novel relationships between superoscillations and modular variables *J. Phys. Conf. Series vol 70* 012016
[52] Tollaksen J 2007 Quantum properties that are extended in time *Quantum Information and Computation V (Proceedings of SPIE vol 6573)* ed E Donkor, A Pirich and H Brandt (Bellingham, WA: SPIE Optical Engineering Press) CID 6573-35
[53] Tollaksen J and Ghoshal D 2007 Weak measurements, weak values and entanglement *Quantum Information and Computation V (Proceedings of SPIE vol 6573)* ed E Donkor, A Pirich and H Brandt (Bellingham, WA: SPIE Optical Engineering Press) CID 6573-36
[54] Tollaksen J 2007 Non-statistical weak measurements *Quantum Information and Computation V (Proceedings of SPIE vol 6573)* ed E Donkor, A Pirich and H Brandt (Bellingham, WA: SPIE Optical Engineering Press) CID 6573-33
[55] Parks A D, Cullin D W and Stoudt D C 1979 *Proc. R. Soc.* A **454**
Parks A D, Cullin D W and Stoudt D C 1998 *Proc. R. Soc.* A **454** 2997–3008
[56] Pryde G J, O'Brien J L, White A G, Ralph T C and Wiseman H M 2005 *Phys. Rev. Lett.* **94** 220405
[57] Ritchie N W M, Story J G and Hulet R G 1991 *Phys. Rev. Lett.* **66** 1107
[58] Ahnert S E and Payne M C 2004 *Phys. Rev.* A **70** 042102
[59] Wiseman H M 2002 *Phys. Rev.* A **65** 032111