Causality Relations and Hidden Variable Theories for the Mermin-Peres Square

Hollis Williams*
*School of Engineering, University of Warwick, Coventry CV4 7AL, UK

It is well-known that eigenvalues in quantum mechanics cannot be assigned to physical properties independently of the measurement context. We argue that it might be possible to relate the contexts of the Mermin-Peres magic square using causality relations in a way which makes explanations using hidden variable models unnecessary and unappealing.

I. INTRODUCTION

It is commonly believed that the advent of quantum mechanics overthrew classical mechanics, but identifying exactly what it is about the experimental predictions of quantum mechanics which cannot be recovered in some sense by classical theories has turned out to be a notoriously thorny issue. A debate has raged for many decades as to whether it is possible for quantum mechanical phenomena to be explained by classical hidden variable theories which introduce unobserved deterministic hidden realities. The Kochen-Specker theorem is a foundational result which states that noncontextual hidden variable theories cannot explain the fundamental relations between physical properties [1]. This in turn implies that quantum mechanics can be interpreted instead as a contextual hidden variable theory. There have now been many related proofs of the Kochen-Specker theorem and no hidden variables, including Greenberger-Horne-Zeilinger–type proofs, a state-independent proof and observable-based geometric and graph theoretic proofs [2-4]. However, the ultimate significance of the theorem is still unclear, especially as the notion of contextuality has never been given a fully satisfactory explanation. The issue is not helped much by the fact that noncontextual hidden variable models work for some problems [5, 6]. In fact, it can be shown that given a certain communication trade-off, classical local hidden variable theories can always be augmented to simulate quantum entangled systems [7].

More generally, contextuality (or noncontextuality) is a property of any system of random variables, where each variable is labelled by its content (the property being measured) and its context (the set of circumstances under which the measurement is recorded, which includes the other random variables with which it is recorded). There are several possible notions of contextuality and measures of contextuality, including the contextual fraction, which can take values between 0 and 1 [8]. Such a measure has to take into account the fact that a context can have a measurement-dependent part and a state preparation part. Weak measurements can recover contextual values because their measurement interaction is too weak to establish its own context. On the computational side, it has been shown that various notions of contextuality are necessary resources for a range of quantum procedures (including a class of quantum computation schemes on qubits) [9-11].

The main issue with hidden variables on the physical side is their apparently ad hoc introduction into the theory to rescue probabilistic measurement outcomes. It has recently been argued that quantum uncertainties differ from classical fluctuations because microscopic fluctuations include quantum coherence [12]. Context is established by some particular combination of state preparation dynamics and measurement, where the state preparation brings quantum fluctuations into the system and the measurement then samples the fluctuations in a specific quantum coherent way. For this reason, it would obviously be very interesting if there were a more physical explanation as to why properties of a system depend on the context. In fact, it is not immediately clear that the Kocken-Specker theorem and related no-go theorems actually have a physical interpretation without setting some further requirements [13]. This would perhaps be more satisfying than the above logical and mathematical proofs, since it would show that hidden variable theories are somehow aesthetically unappealing or even unnecessary, and that they can be cut away in some philosophical sense. This is only a statement of principle, however, since as stated earlier hidden variable models can often be useful depending on the problem in question.

The discussion of quantum fluctuations and contextuality is clearly linked to the measurement problem, since several measurement theories in the literature suggest that these fluctuations represent a dynamical randomness [14, 15]. Quantum contextuality enters here because by contextuality the different possible results which we observe in different measurements come from the same randomness, but they appear as different sets of possibilities depending on the measurement dynamics. Since it is impossible to measure or observe a system without using a coherent context, it follows by this type of argument that contextuality cannot be separated from the measurement problem.

Discussions of contextuality and noncontextuality are intimately connected with the Kochen-Specker theorem. However, there is a problem with the usual claim that the Kochen-Specker theorem removes the possibility of explaining quantum mechanical systems with noncontex-

* Hollis.Williams@warwick.ac.uk

tual hidden variable models, since the standard proofs of the theorem are not valid if one allows the post-measurement probability distribution for the hidden variables to depend on the choice of the set of mutually commuting observables chosen for the measurement [16, 17]. Rather than looking for an observable-based proof of the Kochen-Specker theorem, we will here try to argue that one can use causality relations to relate the measurement contexts for a particular arrangement of operators called the Mermin-Peres square such that hidden variable models are unnecessary.

II. THE MERMIN-PERES SQUARE

We start with the four-dimensional Hilbert space for a pair of spin-1/2 particles. Any observable for a single spin-1/2 particle can be written as a linear combination of the elements of the Pauli group. The Mermin-Peres magic square consists of nine operators arranged in a square matrix as

$$\begin{bmatrix} \hat{\sigma}_x \otimes \mathbb{1} & \mathbb{1} \otimes \hat{\sigma}_x & \hat{\sigma}_x \otimes \hat{\sigma}_x \\ \mathbb{1} \otimes \hat{\sigma}_y & \hat{\sigma}_y \otimes \mathbb{1} & \hat{\sigma}_y \otimes \hat{\sigma}_y \\ \hat{\sigma}_x \otimes \hat{\sigma}_y & \hat{\sigma}_y \otimes \hat{\sigma}_x & \hat{\sigma}_z \otimes \hat{\sigma}_z \end{bmatrix}, \tag{1}$$

where $\hat{\sigma}_i$ denote the usual Pauli matrices [18, 19]. We will denote by $\hat{V}_{ij}$ the operator in row i and column j. The operators in each row and each column are mutually commuting and can therefore be measured jointly. Note that the product of the three operators $\hat{V}_{11}\hat{V}_{12}\hat{V}_{13} = \hat{V}_{21}\hat{V}_{22}\hat{V}_{23} = \hat{V}_{31}\hat{V}_{32}\hat{V}_{33} = +\mathbb{1}\otimes\mathbb{1}$, as does $\hat{V}_{11}\hat{V}_{21}\hat{V}_{31}$ and $\hat{V}_{12}\hat{V}_{22}\hat{V}_{32}$. However, $\hat{V}_{13}\hat{V}_{23}\hat{V}_{33} = -\mathbb{1}\otimes\mathbb{1}$. The operators which sit in the same row or the same column commute, but an operator anti-commutes with an operator which is not in its row or column. Contextual hidden variable models have been constructed for the magic square [13]. It has also been shown that a noncontexual hidden variable theory is possible for the Mermin-Peres square if one allows hidden variable states that can be disturbed via measurement, hence the pre-measurement value of an observable can be different from its post-measurement outcome [16].

Assume that the values of the product operators are not changed by changing the context, in which case it should not matter if the an operator is measured in a row of the magic square rather than a column. If we multiply the products of the three operators in each row, we obtain

$$(\hat{V}_{11}\hat{V}_{12}\hat{V}_{13})(\hat{V}_{21}\hat{V}_{22}\hat{V}_{23})(\hat{V}_{31}\hat{V}_{32}\hat{V}_{33}) = +1. \tag{2}$$

On the other hand, if we multiply the products of the three operators in each column, we obtain

$$(\hat{V}_{11}\hat{V}_{21}\hat{V}_{31})(\hat{V}_{12}\hat{V}_{22}\hat{V}_{32})(\hat{V}_{13}\hat{V}_{23}\hat{V}_{33}) = -1. \tag{3}$$

This would appear to be something of a paradox given that the same operators are involved in both cases, so the values of the product operators must somehow depend on whether they appear in a row or a column of the square.

In [16], similar relations are written down for a collection of functions $V_{ij} : \Omega \to \mathbb{R}$ such that the outcome of measuring the operator $\hat{V}_{ij}$ is $V_{ij}(\omega)$, where $\omega \in \Omega$ and Ω is a set of hidden variables. If Ω and all the V_{ij} are known, the operator relations for products of rows and columns suggest a similar relation in the hidden variables model. Although it is clearly possible to give the magic square a formal interpretation as a noncontextual hidden variables model, it is perhaps less obvious why this interpretation should be necessary or desirable. There are of course other arguments in the literature which try to show that hidden variable theories are unphysical, but our suggestion is somewhat different and we will instead argue that hidden variable theories are in some sense unnecessary or superfluous [20, 21].

We will start by setting down some notation. Let $\mathcal{S}$ denote the set of observables, where A_i denotes an observable. In our case, there are nine observables so $\mathcal{S} = \{A_1, ..., A_9\}$. A context S_α is taken to be a subset of S composed only of mutally commuting operators, where α labels different contexts. The α-th context is written as $S_\alpha = \{A_{k_1^\alpha}, ..., A_{k_m^\alpha}\}$, where the subscript takes values between 1 and 9 and m is the number of observables in the context. In essence, contexts are defined as differences amongst procedures that are operationally equivalent, so it is possible to have a theory in which a notion of noncontextuality can be defined although there is no notion of context [22]. There are a total of six contexts for the Mermin-Peres square:

$$\mathcal{S}_1 = \{A_1, A_2, A_3\}, \quad \mathcal{S}_2 = \{A_4, A_5, A_6\}, \tag{4a}$$

$$\mathcal{S}_3 = \{A_7, A_8, A_9\}, \quad \mathcal{S}_4 = \{A_1, A_4, A_7\}, \tag{4b}$$

$$\mathcal{S}_5 = \{A_2, A_5, A_8\}, \quad \mathcal{S}_6 = \{A_3, A_6, A_9\}. \tag{4c}$$

To match with the notation used in [16], $A_1 = \hat{V}_{11}$, $A_6 = \hat{V}_{23}$, and so on. Note that the simplest Kocken-Specker set which admits a symmetric parity proof has seven contexts [23]. A value assignment v assigns a value to an observable in $A_i \in \mathcal{S}$, where $v(A_i)$ is an eigenvalue of the corresponding operator $\hat{A}_i$. In the Mermin-Peres square, all of the operators square to the identity operator, so $v(A_i) = \pm 1$.

III. CAUSALITY RELATIONS

As stated before, the set of operators in a row or a column (ie. a context) mutually commute in the Mermin-Peres square. As an example, the three operators in $\mathcal{S}_3$ commute, so they share a common eigenstate. One could choose the singlet state $|\psi\rangle$ such that

$$(\sigma_x \otimes \sigma_y)\,|\psi\rangle = -\,|\psi\rangle\,, \tag{5a}$$

$$(\sigma_y \otimes \sigma_x)\,|\psi\rangle = -\,|\psi\rangle\,, \tag{5b}$$

$$(\sigma_z \otimes \sigma_z)\,|\psi\rangle = -\,|\psi\rangle\,. \tag{5c}$$

An initial state can be an eigenstate of $\sigma_x \otimes \sigma_y + \sigma_y \otimes \sigma_x$ without being an eigenstate of each operator individually. Alternatively, the same eigenstate may have a different eigenvalue, so in general the eigenvalue $\nu_{\sigma_x\otimes\sigma_y,\sigma_y\otimes\sigma_x}$ for the sum of two operators will differ from the eigenvalues of the individual operators $\nu_{\sigma_x\otimes\sigma_y}$ and $\nu_{\sigma_y\otimes\sigma_x}$. For example, with the above singlet state, one can have

$$(\sigma_x \otimes \sigma_y + \sigma_y \otimes \sigma_x)\,|\psi\rangle = 0. \tag{6}$$

If we perform a measurement of an eigenvalue of $\sigma_x \otimes \sigma_y$ and determine it to be $+1$ or -1, the value of $\sigma_y \otimes \sigma_x$ depends on the other eigenvalue $\nu_{\sigma_x\otimes\sigma_y}$ and so is written as $\sigma_y \otimes \sigma_x(\nu_{\sigma_x\otimes\sigma_y})$. The value for the eigenvalue of the second operator must now logically be equal to the difference between the two eigenvalues:

$$\sigma_y \otimes \sigma_x(\nu_{\sigma_x\otimes\sigma_y}) = \nu_{\sigma_x\otimes\sigma_y,\sigma_y\otimes\sigma_x} - \nu_{\sigma_x\otimes\sigma_y}\,, \tag{7}$$

where the difference is again equal to ±1 for the singlet state. The same argument also works in reverse, so that we have

$$\sigma_x \otimes \sigma_y(\nu_{\sigma_y\otimes\sigma_x}) = \nu_{\sigma_y\otimes\sigma_x,\sigma_x\otimes\sigma_y} - \nu_{\sigma_y\otimes\sigma_x}\,, \tag{8}$$

Re-arranging and suppressing the arguments, one can then argue that

$$\sigma_x \otimes \sigma_y + \sigma_y \otimes \sigma_x = \nu_{\sigma_y\otimes\sigma_x,\sigma_x\otimes\sigma_y} \tag{9}$$

is effectively a causality relation which relates the possible contexts defined by measuring either of the two commuting operators. In particular, equation (9) relates the context defined by a possible measurement of $\sigma_y \otimes \sigma_x$ (which can either be $\mathcal{S}_3$ or $\mathcal{S}_4$) to the context defined by a measurement of $\sigma_x \otimes \sigma_y$ (which can either be $\mathcal{S}_3$ or $\mathcal{S}_5$).

As is implicit in the example above, the context defined by the measurement of the first operator need not be the same as the context defined by the measurement of the second operator. In the weak values framework, a context is always represented by a certain combination of initial and final conditions [24]. However, in our example above, there is not a definite choice of contexts when measuring both $\sigma_x \otimes \sigma_y$ and $\sigma_y \otimes \sigma_x$ so that the initial singlet state $|\psi\rangle$ defines the causality relation (9) without definitely referring to a context where it is used. When measuring $\sigma_x \otimes \sigma_y$, the usual causality relation is expressed in the eigenstate basis of the other observable $\sigma_y \otimes \sigma_x$ as

$$\left\langle \nu_{\sigma_x\otimes\sigma_y} \,\middle|\, \psi \right\rangle = \sum_{\nu_{\sigma_y\otimes\sigma_x}} \left\langle \nu_{\sigma_x\otimes\sigma_y} \,\middle|\, \nu_{\sigma_y\otimes\sigma_x} \right\rangle \left\langle \nu_{\sigma_y\otimes\sigma_x} \,\middle|\, \psi \right\rangle. \tag{10}$$

This relates the outcome of measuring $\sigma_x \otimes \sigma_y$ to the fluctuations in the initial condition $|\psi\rangle$, where the Hilbert space inner products $\left\langle \nu_{\sigma_y\otimes\sigma_x} \,\middle|\, \psi \right\rangle$ represent the dependence of $\sigma_x \otimes \sigma_y$ on the eigenvalues obtained for $\sigma_y \otimes \sigma_x$.

This suggests that the transition probability amplitude which is written with a summation taken using a basis of eigenstates of $\sigma_x \otimes \sigma_y$ (or oppositely with $\sigma_y \otimes \sigma_x$) represents the same causality relation. For this reason, we argue that a causality relation relates different measurement contexts in a definite way and that hidden variable theories should be effectively unnecessary to explain the value assignments which are observed for the Mermin-Peres square. The causality relations which relate the contexts in $\mathcal{S}$ are reproducible, so it should not be necessary to ponder hidden realities. However, in order to make this a firm conclusion it would be necessary to show that the way that the values of physical properties change when the context is changed is implied by the form of the causality relations. This would require a statistical treatment of measurement contexts and how they emerge from the causality relations between the initial and final conditions [24].

IV. CONCLUSIONS

In conclusion, we have made an attempt to argue that considering reproducible causalities which relate different contexts for the Mermin-Peres magic square might mean that contextual or noncontextual hidden variable theories are not needed. This is especially appealing given the complexity and ad hoc nature of the hidden variable models which have been constructed to explain the measurement results which one sees for the square [16]. This would avoid the absolute necessity of introducing contextual hidden variables into the Mermin-Peres square, either as true hidden variables or as complex weak values which nevertheless assign contextual values to experimental outcomes. It would be an even more difficult question as to whether there is a general explanation for context dependence of physical properties in any scenario which always makes hidden variable theories undesirable.

Note also that we have not used weak values in our argument, although a connection has previously been suggested between weak values and incompatibility of quantum theory with noncontextual hidden variable models [25-29]. It is also worth pointing out that when incorporating weak values into our discussion, it is not necessarily relevant whether those weak values can be observed experimentally, although the literature often concentrates on whether or not these values can be observed. Similar discussions in the literature are phrased in terms of weak values, but we have tried to avoid weak values here because of their still somewhat controversial status (mostly because of the practical difficulties inherent in measuring them) [12]. In summary, it might be possible to avoid invocation of noncontextual and even contexual hidden variable theories to explain the measurement results which can be obtained with a very simple quantum mechanical system (the Mermin-Peres square). The evasion of complex weak values is also desirable, since one still needs somewhat ad hoc constructions to explain how these complex values relate to deterministic real-valued variables which are needed for uncertainty relations [28].

ACKNOWLEDGEMENTS

I would like to thank Holger F. Hofmann for useful discussions and for drawing my attention to reference [24].

REFERENCES

[1] S. Kochen and E. Specker, The Problem of Hidden Variables in Quantum Mechanics, Indiana Univ. Math. J. **17**, 59 (1968).

[2] A. Cabello, Bell's Theorem without Inequalities and without Probabilities for Two Observers, Phys. Rev. Lett. **86**, 1911 (2001).

[3] A. Cabello, "All versus Nothing" Inseparability for Two Observers, Phys. Rev. Lett. **87**, 010403 (2001).

[4] F. Holweck and M. Saniga, Contextuality with a Small Number of Observables, Int. J. Quantum Inform. **15**, 1750026 (2017).

[5] G. Tóth and A. Acín, Genuine tripartite entangled states with a local hidden-variable model, Phys. Rev. A **74**, 030306 (2006).

[6] R.F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[7] G. Brassard, R. Cleve and A. Tapp, Cost of Exactly Simulating Quantum Entanglement with Classical Communication, Phys. Rev. Lett. **83**, 1874 (1999).

[8] S. Abramsky, R.S. Barbosa and S. Mansfield, Contextual Fraction as a Measure of Contextuality, Phys. Rev. Lett. **119**, 050504 (2017).

[9] J. Bermejo Vega, N Delfosse, D.E. Browne, C. Okay and R. Raussendorf, Contextuality as a Resource for Models of Quantum Computation with Qubits, Phys. Rev. Lett. **119**, 120505 (2017).

[10] S. Mansfield and E. Kashefi, Quantum Advantage from Sequential-Transformation Contextuality, Phys. Rev. Lett. **121**, 230401 (2018).

[11] W.M. Kirby and P.J. Love, Contextuality Test of the Nonclassicality of Variational Quantum Eigensolvers, Phys. Rev. Lett. **123**, 200501 (2019).

[12] H.F. Hofmann, Contextuality of quantum fluctuations characterized by conditional weak values of entangled states, Phys. Rev. A **102**, 062215 (2020).

[13] G. Hofer-Szabó, Three noncontextual hidden variable models for the Peres-Mermin square, Eur. J. Phil. Sci. **11**, 30 (2021).

[14] K. Patekar and H.F. Hofmann, The role of system-meter entanglement in controlling the resolution and decoherence of quantum measurements, New. J. Phys. **21**, 103006 (2019).

[15] H. Williams, Superpositions of unitary operators in quantum mechanics, IOP SciNotes **1**, 035204 (2020).

[16] B.R. La Cour, Quantum contextuality in the Mermin-Peres square: A hidden-variable perspective, Phys. Rev. A **79**, 012102 (2009).

[17] X.-D. Yu, Y.-Q. Guo and D.M. Tong, A proof of the Kochen-Specker theorem can always be converted to a state-independent noncontextuality inequality, New J. Phys. **17**, 093001 (2015).

[18] A. Peres, Two simple proofs of the Kochen-Specker theorem, J. Phys. A: Math. Gen. **24**(4), L175 (1991).

[19] N.D. Mermin, Hidden variables and the two theorems of John Bell, Rev. Mod. Phys. **65**(3), 803-815 (1993).

[20] A. Carmi, C. Cohen, L. Maccone and H. Nikolic, Knowledge of Quantum Hidden Variables Enables Backwards-in-Time Signaling, Appl. Sci. **11**, 4477 (2021).

[21] A. Peres, Unperformed experiments have no results, Am. J. Phys **46**, 745 (1978).

[22] D. Schmid, J.H. Selby, E. Wolfe, R. Kunjwal and R.W. Spekkens, Characterization of Noncontextuality in the Framework of Generalized Probabilistic Theories, PRX Quant. **2**, 010331 (2021).

[23] P. Lisoněk, P. Badziag, J Portillo and A Cabello, Kochen-Specker set with seven contexts, Phys. Rev. A **89**, 042101 (2014).

[24] T. Nii, M. Iinuma and H.F. Hofmann, On the relation between measurement outcomes and physical properties, Quantum Stud.: Math. Found. **5**:229-243 (2018).

[25] J. Dressel, S. Agarwal and A.N. Jordan, Contextual Values of Observables in Quantum Measurements, Phys. Rev. Lett. **104**, 240401 (2010).

[26] A. Zeilinger, A foundational principle for quantum mechanics, Found. Phys. **29**, 631 (1999).

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

[28] Agung Budiyono and Hermawan K. Dipojono, Quantum uncertainty as classical uncertainty of real-deterministic variables constructed from complex weak values and a global random variable, Phys. Rev. A **103**, 022215 (2021).

[29] R.W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, Phys. Rev. A **71**, 052108 (2005).