Quantum Advantage from Sequential-Transformation Contextuality

Shane Mansfield[1,*] and Elham Kashefi[1,2]
[1]*Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, F-75005 Paris, France*
[2]*School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, United Kingdom*

(Received 30 January 2018; published 3 December 2018)

We introduce a notion of contextuality for transformations in sequential contexts, distinct from the Bell-Kochen-Specker and Spekkens notions of contextuality. Within a transformation-based model for quantum computation we show that strong sequential-transformation contextuality is necessary and sufficient for deterministic computation of nonlinear functions if classical components are restricted to mod2 linearity and matching constraints apply to any underlying ontology. For probabilistic computation, sequential-transformation contextuality is necessary and sufficient for advantage in this task and the degree of advantage quantifiably relates to the degree of contextuality.

DOI: 10.1103/PhysRevLett.121.230401

Contextuality is a key nonclassical phenomenon exhibited by quantum systems, which was first considered by Bell [1] and by Kochen and Specker [2]. It has been the subject of renewed interest recently as a range of results have established it to be the essential ingredient for enabling quantum advantages over classical implementations of a variety of informatic tasks [3–5], simulation of quantum processes [6], and for enabling universal quantum computing [7–12]. A broader notion of contextuality due to Spekkens *et al.* [16] has also been shown to be essential to quantum advantages relating to state discrimination and one-way communication protocols [17–21]. However, questions remain over which forms of contextuality provide advantage in which precise settings [22] and whether existing notions of contextuality are sufficient to account for all instances of quantum advantage. For example, there exist a variety of advantages achievable with a single qubit [23–25], where Bell-Kochen-Specker (BKS) contextuality cannot arise [2,26] and to which there is no apparent link to the Spekkens version. This raises the important question of which nonclassical feature could be at play if not contextuality of these kinds.

We introduce a notion of contextuality for transformations performed in sequential contexts that is inequivalent to the notion of transformation contextuality introduced by Spekkens [27]. It is necessarily present in a recently discovered form of quantum advantage in shallow circuits [28]. We will show, via a Mermin-style [29,30] parity argument, that it is also crucial in enabling increased computational power in the single qubit example of [25]. The setting for that example is a transformation-based model of quantum computing, which we call here $l2$-TBQC, that was shown to be useful in achieving secure delegated computing. In the model, a classical control computer, whose power is limited to mod2-linear computation, may interact with a quantum resource, by which its computational power may be enhanced. As with the analogous measurement-based model, $l2$-MBQC [4], which was the setting for the results of [3–5], it can provide a useful tool for probing the roots of quantum advantage. In this setting, we show more generally that sequential-transformation contextuality is necessary and sufficient to enable advantage in the task of probabilistically computing any nonlinear function whenever classical ontologies are required to respect the computational assumptions. Moreover, the degree of contextuality can be related to the probability of success, and in particular, strong (i.e., maximal) contextuality is necessary for deterministic computation of any nonlinear function.

Our results trace an arc that parallels developments relating BKS contextuality to quantum advantage in $l2$-MBQC: Anders and Browne provided an example in which a contextual resource is sufficient for the computation of a particular nonlinear function [3]; Raussendorf then proved that strong contextuality is necessary for any deterministic nonlinear computation [4], as initially observed by Hoban *et al.* for nonadaptive $l2$-MBQC [31] based on an early version of [4]; he also showed that contextuality is necessary for quantum advantage in the task of probabilistically computing any nonlinear function; this latter result was later sharpened to show more precisely how the degree of contextuality as measured by the contextual fraction relates to probability of success in [5]. Our results set the stage for further investigation of how sequential-transformation contextuality may relate to quantum advantages, speedups, and the onset of universality in other settings, as the results of [7–10] do for BKS contextuality.

Ontological models.—Quantum theory exhibits a number of apparently nonintuitive features. Crucially, in many cases there exist no-go theorems that establish that there is no way these features can be explained away by recourse to any deeper or more complete theory that would obey certain classical intuitions [32]. Some such nonclassical features are nonlocality [33], (BKS) contextuality [1,2,34], and forms of preparation and transformation contextuality [27], while others relate to macrorealism [43–46] and the ontic nature of the quantum state [47–54]. A convenient formalism for treating such theorems is that of ontological models, which we briefly set out next. Note that in this Letter when we speak of ontological models, we will not be assuming any additional features beyond what is explicitly set out below (e.g., of the kind present in [27]).

The central component is an ontic state space Λ, comprising the states of a hypothetical underlying theory. Preparation of a quantum state ρ results in an ontic state sampled according to a probability distribution d_ρ on Λ [55]. In the simplest case, a quantum transformation U corresponds to a measurable function $f_U\colon \Lambda \to \Lambda$. For consistency we require that $f_{U*}d_\rho = d_{U\rho U^\dagger}$, where the left-hand side is the push forward of d_ρ along f_U, defined by $f_{U*}d_\rho(\lambda) = d_\rho[f_U^{-1}(\lambda)]$. We also require that the function corresponding to the identity operator simply maps each ontic state to the δ function centered on that state, ensuring that $f_{\mathbb{1}*}d_\rho = d_\rho$ for all preparations ρ. In particular, the requirements entail that unitaries correspond to invertible functions. A quantum measurement M corresponds to a function $\xi_M\colon \Lambda \to P(O)$, which assigns to each ontic state a probability distribution over the set of outcomes O. For any combination of preparation, transformation and measurement, the ontological theory predicts that the empirical statistics, $e_{\rho,U,M} \in P(O)$, are given by

$$e_{\rho,U,M} = \sum_{\lambda\in\Lambda} d_\rho(\lambda)\xi_M[f_U(\lambda)]. \tag{1}$$

In fact, our results apply more generally to ontological models in which transformations may correspond to stochastic mixtures of measurable functions. However, we will see shortly that, for our present purposes, since such an ontological model can always be expressed as a convex decomposition of ones in which transformations are deterministic, it will suffice to establish no-go properties for those with deterministic transformations. No-go theorems arise when it is found that ontological models satisfying some additional, perhaps "classical," assumptions are unable to realize the empirical predictions of quantum theory.

(Non-)contextuality.—In the BKS sense, noncontextuality is an assumption of classicality that applies when certain finite sets of compatible measurements may be performed jointly in contexts. It requires that for each valid context C compatibility is reflected at the ontological level through factorizability of the joint measurement function $\xi_C\colon \Lambda \to P(O^{|C|})$; i.e.,

$$\xi_C = \prod_{M\in C} \xi_M. \tag{2}$$

Implicit in this is the crucial requirement that, for any measurement M occurring in contexts C and C', its ontological representation ξ_M is context independent; i.e.,

$$\xi_{M^{(C)}} = \xi_{M^{(C')}}.$$

This description of noncontextuality via factorizability is equivalent to the description in terms of global valuations that may be more familiar to some readers [35].

Next, we mention some specific instances arising from Spekkens' general notion of noncontextuality [27]. Measurement noncontextuality in the explicit sense treated in the no-go results of [27] relaxes (2) to the weaker requirement that

$$\xi_C|_M = \xi_M,$$

for all M and C such that $M \in C$, where $\xi_C|_M$ denotes the marginalization of ξ_C to M.

Transformation and preparation noncontextuality in the explicit sense treated in the no-go results of [27] takes as context any convex decomposition of a given transformation or preparation. This has an operational motivation. Suppose, as a concrete example, that some transformation T admits the following unitary decompositions:

$$T = \frac{1}{2}U_a + \frac{1}{2}U_A \qquad C,$$

$$T = \frac{1}{3}U_a + \frac{1}{3}U_b + \frac{1}{3}U_c. \qquad C'$$

Operationally, context C is "apply U_a or U_A uniformly at random," and context C' is "apply U_a, U_b, or U_c uniformly at random"; quantum mechanically the contexts are equivalent. Noncontextuality requires that convex decompositions are reflected at the ontological level; i.e., in this instance,

$$f_T = \frac{1}{2}f_{U_a} + \frac{1}{2}f_{U_A} = \frac{1}{3}f_{U_a} + \frac{1}{3}f_{U_b} + \frac{1}{3}f_{U_c}.$$

Again, it is implicit that ontological representations of transformations and preparations are independent of operational context; e.g.,

$$f_{U_a^{(C)}} = f_{U_a^{(C')}}.$$

Sequential transformations.—With the preceding versions for comparison, we now introduce a version of noncontextuality for transformations in sequential contexts. It requires that, for each finite sequence of transformations, $C = \{U_i\}_{i=1}^t$, sequential composition is reflected at the ontological level; i.e.,

$$f_{U_t\cdots U_1} = f_{U_t} \circ \ldots \circ f_{U_1}.$$

It is assumed that the ontological representations of transformations are independent of sequential context; i.e., whenever a transformation U occurs in contexts C and C', it holds that

$$f_{U^{(C)}} = f_{U^{(C')}}.$$

When a set of empirical data or predictions cannot be reproduced by an ontological model satisfying this property, it is said to be contextual.

Contextuality in our sense implies that the system of study cannot have an ontology in which transformations correspond to modular, composable operations on ontic states, such that they are well defined independently of which transformations may have been performed previously or will be performed subsequently. Either we must reject the ontological picture entirely or give up on these highly intuitive, classical properties. Note that one plausible, if conspiratorial, mechanism for introducing some contextuality might be through causal dependence on transformations having appeared earlier in the sequence, but even this kind of mechanism is precluded when the transformations being modeled commute.

The constant-depth quantum circuits of [28] provide a concrete example of sequential-transformation contextuality, as they can at best be simulated by classical circuits whose depth grows logarithmically in the size of the input. If a modular, noncontextual ontological description of gate transformations at each step in the circuit were possible, then it would give rise to classical circuits for the same task, which would also have constant depth. Connections to quantum advantage in this setting will be investigated in future work; here we focus on examples in a more restricted setting.

Quantification.—An empirical model $e = \{e_C\}$, associates with each context C a distribution over observed outcomes [35]. Similar to [5], given any empirical model and appropriate version of contextuality, we may consider convex decompositions of the form

$$e = \omega e^{\mathrm{NC}} + (1-\omega)e', \tag{3}$$

where e^{NC} and e' are also empirical models, and e^{NC} is noncontextual. The maximum value of ω over all such decompositions is the noncontextual fraction of e, written $\mathrm{NCF}(e)$, and correspondingly, the contextual fraction of e is $\mathrm{CF}(e) \coloneqq 1 - \mathrm{NCF}(e)$ [56]. For BKS contextuality, the contextual fraction corresponds to the maximum achievable normalized violation by e of any generalized Bell inequality [5]. Here, however, we use it to quantify sequential-transformation contextuality. Using the terminology of the hierarchy of BKS contextuality introduced in [35], an empirical model is said to be strongly contextual when $\mathrm{CF}(e) = 1$.

For a given experimental scenario, the set of all the possible e^{NC} is convex, and any extremal point corresponds simply to fixing a deterministic function $f_U\colon \Lambda \to \Lambda$ for each transformation U featuring in the scenario. Strong contextuality thus arises in the extreme case that no global assignment of deterministic functions to transformations is consistent with even a fraction of the empirical behavior.

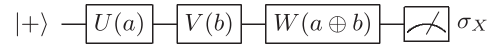

FIG. 1. The basic single qubit AND protocol from [25].

l2-TBQC.—We consider a classical control computer restricted to mod2-linear computation that can interact with a resource, which may be quantum, as follows. The resource is prepared in a fixed state, the control computer may interact with it by means of controlled transformations, then a fixed measurement is performed on the resource, and its outcome returned to the control computer. This captures, for example, the single qubit protocols of [25], which were considered for their security features in a setting in which a client delegates certain operations making up an $l2$-TBQC, such as state preparation and measurement, to a server.

Note that, independent of the $l2$ restriction, any measurement-based quantum computation [61] can equivalently be expressed as a TBQC, since choice of a measurement setting is equivalent to choice of a transformation prior to a fixed measurement.

An example of an $l2$-TBQC that performs a basic nonlinear function, the AND gate on classical input bits a and b,

$$g(a,b) = (a \oplus 1) \otimes (b \oplus 1) \oplus 1,$$

is the following (Fig. 1) [62]. The control computer receives inputs a and b. For the resource, the fixed state is the qubit state $|+\rangle$, the fixed measurement is given by the Pauli operator σ_X, and the controlled transformations are $U(a)$, then $V(b)$, then $W(a \oplus b)$, where

$$U(0) = V(0) = W(0) = I,$$
$$U(1) = V(1) = W(1) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/2} \end{pmatrix}.$$

Notice that all transformations commute. The output of the computation is the measurement outcome interpreted in $\mathbb{Z}_2$, with eigenvalues $+1$ and -1 mapped to 0 and 1, respectively. In terms of complexity classes, access to a qubit quantum resource promotes the computational power from the class $\oplus L$ [63,64] to P, as with the example in [3] in the setting of $l2$-MBQC.

$\oplus L$ ontology.—Of course, classical computers can perfectly well compute nonlinear functions and they also constitute valid noncontextual ontologies. To pose a meaningful computational question about whether a resource may be used to boost power from $\oplus L$ to P, therefore, we will restrict attention to $\oplus L$ ontologies, which we define as follows. Recalling that $\oplus L$ circuits are built entirely of NOT and controlled-NOT (CNOT) gates [64], we will suppose that available transformations are built from these and act on an ontic state space $\mathbb{Z}_2^s$, for some $s \in \mathbb{N}$. In what follows, we will be interested in protocols in which transformations commute. These can already permit efficient solutions to problems for which it is believed there can be no efficient

classical solution [65]. For transformations in commutative $\oplus L$ ontologies, it holds that, for any transformation U,

$$f_U(\boldsymbol{\lambda}) = (I \oplus A_U)\boldsymbol{\lambda} \oplus \boldsymbol{u}, \tag{4}$$

where A_U is an $s \times s$ matrix over $\mathbb{Z}_2$ containing only off-diagonal entries and $\boldsymbol{u} \in \mathbb{Z}_2^s$ (see Supplemental Material [66]). For composition of transformations $\{U_i\}_{i=1}^t$ with ontological representations determined by $\{A_i, \boldsymbol{u}_i\}$, it holds that

$$f_{U_t} \circ \cdots \circ f_{U_1}(\boldsymbol{\lambda}) = \boldsymbol{\lambda} \oplus \bigoplus_{i=1}^{t} A_i \boldsymbol{\lambda} \oplus \bigoplus_{i=1}^{t} \boldsymbol{u}_i.$$

A dichotomic measurement in an $\oplus$ L ontology can most generally be described by a transformation followed by output of the bit value of a fixed entry j of the final ontic state vector; i.e., $\boldsymbol{\lambda}' \cdot \boldsymbol{\delta}$ where $\boldsymbol{\lambda}', \boldsymbol{\delta} \in \mathbb{Z}_2^s$ are the post-transformation ontic state and the vector with jth entry 1 and 0's elsewhere, respectively.

Proposition.—Any commutative $\oplus L$-ontological realization of the AND $l2$-TBQC is transformation contextual.

Proof.—Suppose that preparation results in an initial ontic state $\boldsymbol{\lambda} \in (\mathbb{Z}_2)^s$. From (1), noncontextual realization of the protocol requires Eqs. (5)–(8) to be satisfied. These describe evaluation of the computation for the four possible sequential contexts, where ontological representations of $U(k)$, $V(k)$, and $W(k)$, with $k \in \{0,1\}$, are determined through Eq. (4) by $\{A_U(k), \boldsymbol{u}(k)\}, \{A_V(k), \boldsymbol{v}(k)\}$, and $\{A_W(k), \boldsymbol{w}(k)\}$, respectively, and of the transformation component of the measurement by $\{A_M, \boldsymbol{m}\}$,

$$[\boldsymbol{\lambda} \oplus A_U(0)\boldsymbol{\lambda} \oplus A_V(0)\boldsymbol{\lambda} \oplus A_W(0)\boldsymbol{\lambda} \oplus A_M\boldsymbol{\lambda} \oplus \boldsymbol{u}(0) \oplus \boldsymbol{v}(0) \oplus \boldsymbol{w}(0) \oplus \boldsymbol{m}] \cdot \boldsymbol{\delta} = 0, \tag{5}$$

$$[\boldsymbol{\lambda} \oplus A_U(0)\boldsymbol{\lambda} \oplus A_V(1)\boldsymbol{\lambda} \oplus A_W(1)\boldsymbol{\lambda} \oplus A_M\boldsymbol{\lambda} \oplus \boldsymbol{u}(0) \oplus \boldsymbol{v}(1) \oplus \boldsymbol{w}(1) \oplus \boldsymbol{m}] \cdot \boldsymbol{\delta} = 0, \tag{6}$$

$$[\boldsymbol{\lambda} \oplus A_U(1)\boldsymbol{\lambda} \oplus A_V(0)\boldsymbol{\lambda} \oplus A_W(1)\boldsymbol{\lambda} \oplus A_M\boldsymbol{\lambda} \oplus \boldsymbol{u}(1) \oplus \boldsymbol{v}(0) \oplus \boldsymbol{w}(1) \oplus \boldsymbol{m}] \cdot \boldsymbol{\delta} = 0, \tag{7}$$

$$[\boldsymbol{\lambda} \oplus A_U(1)\boldsymbol{\lambda} \oplus A_V(1)\boldsymbol{\lambda} \oplus A_W(0)\boldsymbol{\lambda} \oplus A_M\boldsymbol{\lambda} \oplus \boldsymbol{u}(1) \oplus \boldsymbol{v}(1) \oplus \boldsymbol{w}(0) \oplus \boldsymbol{m}] \cdot \boldsymbol{\delta} = 1. \tag{8}$$

Under the assumption of noncontextuality, the equations are not jointly satisfiable. This can be deduced from the fact that the sum modulo 2 of the right-hand sides is one, whereas the sum of the left-hand sides is zero, since each vector appears an even number of times, leading to cancellations. Note that a contextual realization would permit ontological representations to vary according to context; e.g., $\boldsymbol{u}(0)^{(4)} \neq \boldsymbol{u}(0)^{(5)}$. Contextually, we can always satisfy the equations. The conclusion is that, while $\oplus L$-ontological descriptions are possible, they are necessarily transformation contextual. ■

The above proof is similar to Mermin's parity version [29,30] of the Greenberger-Horne-Shimony-Zeilinger inequality-free argument for nonlocality [67,68] and is an instance of an all-versus-nothing proof of strong contextuality [69], albeit for transformation rather than BKS contextuality.

Proposition.—Strong transformation contextuality is necessary for $\oplus L$-ontological realization of any nonlinear commutative $l2$-TBQC.

The proof is in the Supplemental Material [66].

Given two functions $g, h\colon \mathbb{Z}_2^r \to \mathbb{Z}_2$, we can define an average distance between these functions as

$$d(g,h) \coloneqq 2^{-r}|\{\boldsymbol{i}|g(\boldsymbol{i}) \neq h(\boldsymbol{i})\}|.$$

This can be used to measure the degree of nonlinearity of any function $g\colon \mathbb{Z}_2^r \to \mathbb{Z}_2$ as the distance to the closest linear function of that type,

$$\nu(g) \coloneqq \min\{d(g,h)|h\colon \mathbb{Z}_2^r \to \mathbb{Z}_2 \text{linear}\}.$$

Theorem.—If a commutative $l2$-TBQC, with resource empirical model e, probabilistically computes a function $g\colon \mathbb{Z}_2^r \to \mathbb{Z}_2$ with an average failure probability ε over all 2^r possible inputs, then

$$\varepsilon \geq \mathrm{NCF}(e)\nu(g).$$

The theorem extends the preceding proposition, since, in particular, it implies that deterministic computation ($\varepsilon = 0$) of a nonlinear function [$\nu(g) > 0$] requires strong contextuality [$\mathrm{NCF}(e) = 0$]. The proof is in the Supplemental Material [66] and is similar to that of Theorem 3 in [5].

Discussion.—The present results highlight the potential of sequential contextuality as a source of quantum advantage of a single qubit over arbitrarily many classical bits for a particular kind of computational task. While the $\oplus L$-ontological assumptions are natural in the particular setting of restricted classical computation that we consider, a direction for future research will be to consider examples of sequential-transformation contextuality in less restricted settings, like that of [28], as well as to explore other potential connections to quantum advantage, especially in single qubit systems [23,24]. It also remains to be seen how the present notion of contextuality can be treated in resource-theoretic frameworks of the kind developed in [5,57,70–72]. A related analysis, in terms of irreversibility, of transformation-based protocols is contained in [73], and in the future, it may be interesting to consider advantages as arising from a combination of these phenomena [74]. From a foundational perspective, in light of the present analysis, the experimental results of [75,76] could already be said to provide indirect experimental evidence for a kind of sequential-transformation contextuality, but this leaves open the possibility for experiments designed specifically to test for the feature, which might also aim to minimize potential issues, such as the detection loophole.

The authors thank Samson Abramsky, Rui Soares Barbosa, Dan Browne, Ulysse Chabaud, Tom Douce, Pierre-Emmanuel Emeriau, Ernesto Galvão, Frédéric Grosshans, Luciana Henaut, Matty Hoban, Aleks Kissinger, Jan-Åke Larsson, Damian Markham, and Anna Pappa for valuable discussions and comments. Early ideas for this work were conceived while S. M. was visiting the Simons Institute for the Theory of Computing at the University of California, Berkeley, as a participant of the Logical Structures in Computation program. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 750523.

*shane.mansfield@lip6.fr

[1] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[2] S. Kochen and E. P. Specker, in *The Logico-Algebraic Approach to Quantum Mechanics* (Springer Netherlands, New York, 1975), pp. 263–276.
[3] J. Anders and D. E. Browne, Phys. Rev. Lett. **102**, 050502 (2009).
[4] R. Raussendorf, Phys. Rev. A **88**, 022322 (2013).
[5] S. Abramsky, R. S. Barbosa, and S. Mansfield, Phys. Rev. Lett. **119**, 050504 (2017).
[6] A. Karanjai, J. J. Wallman, and S. D. Bartlett, arXiv:1802.07744.
[7] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature **510**, 351 (2014).
[8] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf, Phys. Rev. X **5**, 021003 (2015).
[9] H. Pashayan, J. J. Wallman, and S. D. Bartlett, Phys. Rev. Lett. **115**, 070501 (2015).
[10] J. Bermejo-Vega, N. Delfosse, D. E. Browne, C. Okay, and R. Raussendorf, Phys. Rev. Lett. **119**, 120505 (2017).
[11] L. Catani and D. E. Browne, arXiv:1711.08676 [Phys. Rev. A (to be published)].
[12] Some of these references build on earlier work relating quantum advantage to Wigner function negativity [13,14], known to be an equivalent notion of nonclassicality to contextuality [15].
[13] E. F. Galvão, Phys. Rev. A **71**, 042302 (2005).
[14] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New J. Phys. **14**, 113011 (2012).
[15] R. W. Spekkens, Phys. Rev. Lett. **101**, 020401 (2008).
[16] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Phys. Rev. Lett. **102**, 010401 (2009).
[17] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, New J. Phys. **18**, 045003 (2016).
[18] D. Saha, P. Horodecki, and M. Pawłowski, arXiv:1708.04751.
[19] D. Schmid and R. W. Spekkens, Phys. Rev. X **8**, 011015 (2018).
[20] D. Saha and A. Chaturvedi, arXiv:1802.07215.
[21] S. Ghorai and A. K. Pan, arxiv:1806.01194.
[22] P. Lillystone, J. J. Wallman, and J. Emerson, arXiv:1802.06121.
[23] E. Knill and R. Laflamme, Phys. Rev. Lett. **81**, 5672 (1998).
[24] E. F. Galvão and L. Hardy, Phys. Rev. Lett. **90**, 087902 (2003).
[25] V. Dunjko, T. Kapourniotis, and E. Kashefi, Quantum Inf. Comput. **16**, 0061 (2016).
[26] A. M. Gleason, J. Math. Mech. **6**, 885 (1957).
[27] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[28] S. Bravyi, D. Gosset, and R. Koenig, Science **362**, 308 (2018).
[29] N. D. Mermin, Am. J. Phys. **58**, 731 (1990).
[30] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).
[31] M. J. Hoban, E. T. Campbell, K. Loukopoulos, and D. E. Browne, New J. Phys. **13**, 023014 (2011).
[32] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[33] J. S. Bell, Physics **1**, 195 (1964).
[34] There now exist a number of unified treatments of nonlocality and (BKS) contextuality [35–38], which are closely interrelated [36,39], and whose relations to the other kinds of nonclassicality have also been explored elsewhere [40–42].
[35] S. Abramsky and A. Brandenburger, New J. Phys. **13**, 113036 (2011).
[36] A. Acín, T. Fritz, A. Leverrier, and A. B. Sainz, Commun. Math. Phys. **334**, 533 (2015).
[37] A. Cabello, S. Severini, and A. Winter, Phys. Rev. Lett. **112**, 040401 (2014).
[38] E. N. Dzhafarov, J. V. Kujala, and V. H. Cervantes, in *International Symposium on Quantum Interaction* (Springer, New York, 2015), pp. 12–23.
[39] N. de Silva, Phys. Rev. A **95**, 032108 (2017).
[40] J. V. Kujala, E. N. Dzhafarov, and J.-Å. Larsson, Phys. Rev. Lett. **115**, 150401 (2015).
[41] L. Wester, in *Proceedings of 14th International Conference on Quantum Physics and Logic, Nijmegen, The Netherlands, 2017, Electronic Proceedings in Theoretical Computer Science*, Vol. 266, edited by B. Coecke and A. Kissinger (Open Publishing Association, Waterloo, 2018), pp. 1–22.
[42] R. Kunjwal, arXiv:1709.01098.
[43] A. J. Leggett and A. Garg, Phys. Rev. Lett. **54**, 857 (1985).
[44] C. Brukner, S. Taylor, S. Cheung, and V. Vedral, arXiv:quant-ph/0402127.
[45] C. Timpson and O. Maroney, arXiv:1412.6139.
[46] J.-M. A. Allen, O. J. E. Maroney, and S. Gogioso, Quantum **1**, 13 (2017).
[47] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Phys. Rep. **506**, 1 (2011).
[48] M. F. Pusey, J. Barrett, and T. Rudolph, Nat. Phys. **8**, 475 (2012).
[49] R. Colbeck and R. Renner, Phys. Rev. Lett. **108**, 150402 (2012).
[50] R. Colbeck and R. Renner, New J. Phys. **19**, 013016 (2017).

[51] L. Hardy, Int. J. Mod. Phys. B **27**, 1345012 (2013).
[52] A. Montina, Mod. Phys. Lett. A **30**, 1530001 (2015).
[53] S. Mansfield, Phys. Rev. A **94**, 042124 (2016).
[54] J.-M. A. Allen, Quantum Stud. Math. Found. **3**, 161 (2016).
[55] Ontological theories can be defined more generally with measures, but this will not be necessary for our present purposes.
[56] BKS contextuality has also been quantified in such a manner in Refs. [35,57–60].
[57] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik, Phys. Rev. Lett. **112**, 120401 (2014).
[58] E. Amselem, L. E. Danielsen, A. J. López-Tarrida, J. R. Portillo, M. Bourennane, and A. Cabello, Phys. Rev. Lett. **108**, 200405 (2012).
[59] S. Mansfield, D. Phil. thesis, Oxford University, 2013.
[60] S. Abramsky, R. S. Barbosa, and S. Mansfield, in *Informal Proceedings of Quantum Physics & Logic* (2016).
[61] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).
[62] This is a slightly simplified version of the implementation from Ref. [25]. Note that in the homomorphism from $\mathbb{Z}_2$ to the Booleans, which one would perform in order to interpret the function as the logical AND gate, the roles of zero and one are exchanged, i.e., $0 \mapsto 1$ and $1 \mapsto 0$.
[63] C. Damm, Inf. Proc. Lett. **36**, 247 (1990).
[64] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).
[65] M. J. Bremner, R. Jozsa, and D. J. Shepherd, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (The Royal Society, London, 2010), p. rspa20100301.
[66] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.121.230401 for a brief comparison of the present notion of contextuality with that of Spekkens, sufficient conditions for commutativity in $\oplus L$ ontologies, and additional mathematical proofs.
[67] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory and Conceptions of the Universe* (Springer, Dordrecht, 1989), pp. 69–72.
[68] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990).
[69] S. Abramsky, R. S. Barbosa, K. Kishida, R. Lal, and S. Mansfield, in *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*, edited by S. Kreutzer, Leibniz International Proceedings in Informatics (LIPIcs) Vol. 41 (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, 2015), pp. 211–228.
[70] K. Horodecki, A. Grudka, P. Joshi, W. Kłobus, and J. Łodyga, Phys. Rev. A **92**, 032104 (2015).
[71] B. Amaral, A. Cabello, M. T. Cunha, and L. Aolita, Phys. Rev. Lett. **120**, 130403 (2018).
[72] C. Duarte and B. Amaral, J. Math. Phys. **59**, 062202 (2018).
[73] L. Henaut, L. Catani, D. E. Browne, S. Mansfield, and A. Pappa, arXiv:1806.05624 [Phys. Rev. A (to be published)].
[74] L. Catani, Ph.D. thesis, University College London, 2018.
[75] S. Barz, V. Dunjko, F. Schlederer, M. Moore, E. Kashefi, and I. A. Walmsley, Phys. Rev. A **93**, 032339 (2016).
[76] M. Clementi, A. Pappa, A. Eckstein, I. A. Walmsley, E. Kashefi, and S. Barz, Phys. Rev. A **96**, 062317 (2017).