Contextual Fraction as a Measure of Contextuality

Samson Abramsky,[1] Rui Soares Barbosa,[1] and Shane Mansfield[2]
[1]Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom
[2]School of Informatics, University of Edinburgh, Informatics Forum, 10 Crichton Street, Edinburgh EH8 9AB, United Kingdom

(Received 2 February 2017; published 4 August 2017)

We consider the contextual fraction as a quantitative measure of contextuality of empirical models, i.e., tables of probabilities of measurement outcomes in an experimental scenario. It provides a general way to compare the degree of contextuality across measurement scenarios; it bears a precise relationship to violations of Bell inequalities; its value, and a witnessing inequality, can be computed using linear programing; it is monotonic with respect to the "free" operations of a resource theory for contextuality; and it measures quantifiable advantages in informatic tasks, such as games and a form of measurement-based quantum computing.

DOI: 10.1103/PhysRevLett.119.050504

Introduction.—Recent results have established the role of contextuality as a resource for increasing the computational power of specific models of computation [1,2], including enabling universal quantum computation [3]. From this perspective, it is particularly relevant to look for appropriate measures of contextuality and indeed to pose the question of what constitutes a good measure.

In this Letter, we propose a measure of contextuality—the contextual fraction—which provides a quantitative grading between noncontextuality, at one extreme, and maximal contextuality, at the other. A maximally contextual empirical model is one that admits no proper decomposition into a convex combination of a noncontextual model and another model. In this sense, it is meaningful to consider both the noncontextual and contextual fractions of any no-signaling empirical model.

These definitions are made in the general setting of the approach to contextuality introduced in Ref. [4], in which nonlocality is seen as a special case of contextuality.

We show that the contextual fraction has a number of desirable properties: (i) It is fully general in the sense that it applies in any measurement scenario; (ii) it is bounded and normalized, taking values in the interval [0, 1], with 0 indicating noncontextuality and 1 indicating strong contextuality, so it may be used to sensibly compare the degree of contextuality of empirical models not just in a given measurement scenario but also across scenarios; (iii) it has a precise relationship with violations of Bell inequalities, being the maximum normalized violation attained by the empirical model for any Bell inequality on the corresponding measurement scenario; (iv) both the contextual fraction and a witnessing Bell inequality are computable using linear programing—these were implemented and used for computational exploration of some quantum examples; (v) it is monotonic with respect to a range of operations on empirical models that intuitively do not generate contextuality and thus constitute natural "free" operations in a resource theory of contextuality, analogous to the resource theory of entanglement under local operations and classical communication [5], and subsuming existing resource theories for nonlocality [6–9]; (vi) finally, it is related to a quantifiable increase of computational power in a certain form of measurement-based quantum computation, sharpening the results of Ref. [2], and similarly to advantage in games.

We leave for future work an analysis of the relationship between the contextual fraction and other possible measures [11] and further development of (vi).

General framework for contextuality.—We briefly summarize the framework introduced in Ref. [4]. The main objects of study are *empirical models*: tables of data, specifying probability distributions over the joint outcomes of sets of compatible measurements.

A *measurement scenario* is an abstract description of a particular experimental setup. It consists of a triple $\langle X, \mathcal{M}, O\rangle$, where X is a finite set of measurements, O is a finite set of outcome values for each measurement, and $\mathcal{M}$ is a set of subsets of X. Each $C \in \mathcal{M}$ is called a *measurement context* and represents a set of measurements that can be performed together.

Examples of measurement scenarios include multipartite Bell-type scenarios familiar from discussions of nonlocality, Kochen-Specker configurations, measurement scenarios associated with qudit stabilizer quantum mechanics, and more. For example, the well-known (2, 2, 2) Bell scenario, where two experimenters, Alice and Bob, can each choose between performing one of two different measurements, say, a_1 or a_2 for Alice and b_1 or b_2 for Bob, obtaining one of two possible outcomes, is represented as follows:

$$X = \{a_1, a_2, b_1, b_2\}, \qquad O = \{0, 1\},$$
$$\mathcal{M} = \{\{a_1, b_1\}, \{a_1, b_2\}, \{a_2, b_1\}, \{a_2, b_2\}\}.$$

Given this description of the experimental setup, then either performing repeated runs of such experiments with varying choices of measurement context and recording the frequencies of the various outcome events, or calculating theoretical predictions for the probabilities of these outcomes, results in a probability table as in Table I.

TABLE I. Two empirical models on the (2, 2, 2) Bell scenario: the well-known Clauser-Horne-Shimony-Holt model [20], obtained from local projective measurements equatorial at angles 0 (for a_1, b_1) and $\pi/3$ (for a_2, b_2) on the maximally entangled two-qubit Bell state $|\Phi^+\rangle = (1/\sqrt{2})(|00\rangle + |11\rangle)$, and the Popescu-Rohrlich box [21].

A	B	00	01	10	11
a_1	b_1	1/2	0	0	1/2
a_1	b_2	3/8	1/8	1/8	3/8
a_2	b_1	3/8	1/8	1/8	3/8
a_2	b_2	1/8	3/8	3/8	1/8

A	B	00	01	10	11
a_1	b_1	1/2	0	0	1/2
a_1	b_2	1/2	0	0	1/2
a_2	b_1	1/2	0	0	1/2
a_2	b_2	0	1/2	1/2	0

Such data are formalized as an empirical model for the given measurement scenario $\langle X, \mathcal{M}, O\rangle$. For each measurement context C, there is a probability distribution e_C on the joint outcomes of performing all the measurements in C, that is, on the set O^C of functions assigning an outcome in O to each measurement in C.

We require that the marginals of these distributions agree whenever contexts overlap, i.e.,

$$\forall C, C' \in \mathcal{M}, \qquad e_C|_{C\cap C'} = e_{C'}|_{C\cap C'},$$

where the notation $e_C|_U$ with $U \subseteq C$ stands for marginalization of probability distributions (to "forget" the outcomes of some measurements): For $t \in O^U$, $e_C|_U(t) \coloneqq \sum_{s\in O^C, s|_U=t} e_C(s)$. The requirement of *compatibility of marginals* is a generalization of the usual no-signaling condition and is satisfied, in particular, by all empirical models arising from quantum predictions [4].

An empirical model is said to be *contextual* if this family of distributions cannot itself be obtained as the marginals of a single probability distribution on global assignments of outcomes to all measurements, i.e., a distribution d on O^X (where O^X acts as a canonical set of deterministic hidden variables) such that $\forall C \in \mathcal{M}$, $d|_C = e_C$. Equivalently [4], contextual empirical models are those which have no realization by factorizable hidden variable models; thus, for Bell-type measurement scenarios, contextuality specializes to the usual notion of *nonlocality.*

In certain cases, one can witness contextuality from merely the possibilistic, rather than probabilistic, information contained in an empirical model—i.e., which events are possible (with nonzero probability) and which are impossible (with zero probability). A yet stronger form of contextuality occurs when no global assignment of outcomes is even consistent with the possible events: An empirical model e is said to be *strongly contextual* if there is no global assignment $g \in O^X$ such that $\forall C \in \mathcal{M}$, $e_C(g|_C) > 0$. An example is the Popescu-Rohrlich box (Table I). This is the highest level in the qualitative hierarchy of strengths of contextuality introduced in Ref. [4].

The contextual fraction.—Given two empirical models e and e' on the same measurement scenario and $\lambda \in [0, 1]$, we define the empirical model $\lambda e + (1-\lambda)e'$ by taking the convex sum of probability distributions at each context. Compatibility is preserved by this convex sum; hence, it yields a well-defined empirical model.

A natural question to ask is, what fraction of a given empirical model e admits a noncontextual explanation? This approach enables a refinement of the binary notion of contextuality vs noncontextuality into a quantitative grading. Instead of asking for a probability distribution on global assignments that marginalizes to the empirical distributions at each context, we ask only for a subprobability distribution [22] b on global assignments O^X that marginalizes at each context to a subdistribution of the empirical data, thus explaining a fraction of the events, i.e., $\forall C \in \mathcal{M}, b|_C \leq e_C$. Equivalently, we ask for a convex decomposition

$$e = \lambda e^{NC} + (1-\lambda)e', \qquad (1)$$

where e^{NC} is a noncontextual model and e' is another (no-signaling) empirical model. The maximum weight of such a global subprobability distribution, or the maximum possible value of λ in such a decomposition, is called the *noncontextual fraction* of e and generalizes the *local fraction* previously introduced for models on Bell-type scenarios [23,24]. We denote it by $\mathsf{NCF}(e)$ and the *contextual fraction* by $\mathsf{CF}(e) \coloneqq 1 - \mathsf{NCF}(e)$.

The notion of contextual fraction in general scenarios was introduced in Ref. [4], where it was proved that a model is strongly contextual if and only if its contextual fraction is 1. In fact, in any convex decomposition of the form (1) giving maximal weight to the noncontextual model, the other model will necessarily be strongly contextual. This means that any empirical model e admits a convex decomposition

$$e = \mathsf{NCF}(e)e^{NC} + \mathsf{CF}(e)e^{SC} \qquad (2)$$

into a noncontextual and a strongly contextual model. Note that e^{NC} and e^{SC} are not necessarily unique.

Computing the contextual fraction via LP.—The task of finding a consistent probability subdistribution with maximum weight for a given empirical model can be formulated as a linear programing problem. This is a relaxation of the test for contextuality by solving a system of linear equations over the non-negative reals from Ref. [4].

Fix a measurement scenario $\langle X, \mathcal{M}, O\rangle$. Let $n \coloneqq |O^X|$ be the number of global assignments, and $m \coloneqq \sum_{C\in\mathcal{M}} |O^C| = |\{\langle C, s\rangle | C \in \mathcal{M}, s \in O^C\}|$ be the number of local assignments ranging over contexts. The *incidence matrix* [4] $\mathbf{M}$ is an $m \times n$ (0,1) matrix that records the restriction relation between global and local assignments:

$$\mathbf{M}[\langle C, s\rangle, g] \coloneqq \begin{cases} 1 & \text{if } g|_C = s; \\ 0 & \text{otherwise.} \end{cases}$$

An empirical model e can be represented as a vector $\mathbf{v}^e \in \mathbb{R}^m$, with the component $\mathbf{v}^e[\langle C, s\rangle]$ recording the probability given by the model to the assignment s at the measurement context C, $e_C(s)$. This vector is a flattened version of the table used to represent the empirical model (e.g., Table I). The columns of the incidence matrix, $\mathbf{M}[-, g]$, are the vectors corresponding to the (noncontextual) deterministic models obtained from global assignments $g \in O^X$. Recall that every noncontextual model can be written as a mixture of these. A probability distribution on global assignments can be represented as a vector $\mathbf{d} \in \mathbb{R}^n$ with non-negative components summing to 1, and then the corresponding noncontextual model is represented by the vector $\mathbf{M}\,\mathbf{d}$. So a model e is noncontextual if and only if there exists a $\mathbf{d} \in \mathbb{R}^n$ such that

$$\mathbf{M}\,\mathbf{d} = \mathbf{v}^e \quad \text{and} \quad \mathbf{d} \geq \mathbf{0}.$$

Note that the first condition implies that $\mathbf{d}$ is normalized.

A global subprobability distribution is also represented by a vector $\mathbf{b} \in \mathbb{R}^n$ with non-negative components, its weight being given by the dot product $\mathbf{1} \cdot \mathbf{b}$, where $\mathbf{1} \in \mathbb{R}^n$ is the vector whose n components are each 1. The following LP thus calculates the noncontextual fraction of an empirical model e, with $\mathsf{NCF}(e) = \mathbf{1} \cdot \mathbf{b}^*$, where $\mathbf{b}^*$ is an optimal solution:

$$\begin{array}{ll} \texttt{Find} & \mathbf{b} \in \mathbb{R}^n \\ \texttt{maximizing} & \mathbf{1} \cdot \mathbf{b} \\ \texttt{subject to} & \mathbf{M}\,\mathbf{b} \leq \mathbf{v}^e \\ \texttt{and} & \mathbf{b} \geq \mathbf{0}. \end{array} \tag{3}$$

Violations of generalized Bell inequalities.—We now provide further justification for viewing the contextual fraction as a measure of contextuality by relating it to violations of contextuality-witnessing inequalities.

An *inequality* for a scenario $\langle X, \mathcal{M}, O\rangle$ is given by a vector $\mathbf{a} \in \mathbb{R}^m$ of real coefficients indexed by local assignments $\langle C, s\rangle$ and a bound R. For a model e, the inequality reads $\mathbf{a} \cdot \mathbf{v}^e \leq R$, where

$$\mathbf{a} \cdot \mathbf{v}^e = \sum_{C \in \mathcal{M}, s \in O^C} \mathbf{a}[\langle C, s\rangle] e_C(s).$$

Without any loss of generality, we can take R to be non-negative (in fact, even $R = 0$) as any inequality is equivalent to one of this form. We call it a *Bell inequality* if it is satisfied by every noncontextual model. This generalizes the usual notion of Bell inequality, which is defined for Bell-type scenarios for nonlocality, to apply to any contextuality scenario. If, moreover, it is saturated by some noncontextual model, the Bell inequality is said to be *tight*. A Bell inequality establishes a bound for the value of $\mathbf{a} \cdot \mathbf{v}^e$ among noncontextual models e. For more general models, this quantity is limited only by the algebraic bound

$$\|\mathbf{a}\| \coloneqq \sum_{C \in \mathcal{M}} \max \{\mathbf{a}[\langle C, s\rangle] | s \in O^C\}.$$

Note that we will consider only inequalities satisfying $R < \|\mathbf{a}\|$, which excludes inequalities trivially satisfied by all models and avoids cluttering the presentation with special caveats about division by 0.

The *violation* of a Bell inequality $\langle \mathbf{a}, R\rangle$ by a model e is $\max\{0, \mathbf{a} \cdot \mathbf{v}^e - R\}$. However, it is useful to normalize this value by the maximum possible violation in order to give a better idea of the extent to which the model violates the inequality. The *normalized violation* of the Bell inequality by the model e is

$$\frac{\max\{0, \mathbf{a} \cdot \mathbf{v}^e - R\}}{\|\mathbf{a}\| - R}.$$

Theorem 1. Let e be an empirical model. (i) The normalized violation by e of any Bell inequality is at most $\mathsf{CF}(e)$; (ii) if $\mathsf{CF}(e) > 0$, this bound is attained; i.e., there exists a Bell inequality whose normalized violation by e is $\mathsf{CF}(e)$; (iii) moreover, for any decomposition of the form (2), this Bell inequality is tight at the noncontextual model e^{NC} [provided $\mathsf{NCF}(e) > 0$] and maximally violated at the strongly contextual model e^{SC}.

The proof of this result is based on the strong duality theorem of linear programing [27]. It provides an LP method of calculating a witnessing Bell inequality for any empirical model e. The symmetric dual of (3) is the following LP:

$$\begin{array}{ll} \texttt{Find} & \mathbf{y} \in \mathbb{R}^m \\ \texttt{minimizing} & \mathbf{y} \cdot \mathbf{v}^e \\ \texttt{subject to} & \mathbf{M}^T \mathbf{y} \geq \mathbf{1} \\ \texttt{and} & \mathbf{y} \geq \mathbf{0}. \end{array} \tag{4}$$

The strong duality theorem says that, if $\mathbf{b}^*$ is a solution for (3), then there is a solution $\mathbf{y}^*$ for (4) satisfying

$$\mathbf{1} \cdot \mathbf{b}^* = \mathbf{y}^* \cdot \mathbf{v}^e. \tag{5}$$

Defining $\mathbf{a}^* \coloneqq |\mathcal{M}|^{-1}\mathbf{1} - \mathbf{y}^*$, one can show using (5) that the Bell inequality determined by $\mathbf{a}^*$ as the vector of coefficients and with bound $R = 0$ satisfies parts (ii) and (iii) of Theorem 1. A detailed proof is provided in Supplemental Material [28].

Monotonicity.—A key desideratum of a useful measure of contextuality is that it be monotonic for the free operations of a resource theory for contextuality. A fuller treatment of this subject will be presented in a forthcoming article by the authors; here, we consider the properties of the contextual fraction with respect to some of these operations.

We consider the following operations: first, translation of measurements (including restriction to a smaller set of measurements, replication of measurements, etc.); second, coarse-graining of outcomes. Special cases of these give isomorphic relabeling of measurements and outcomes. We also consider operations that combine two empirical models to build a new one. The first of these is *probabilistic*

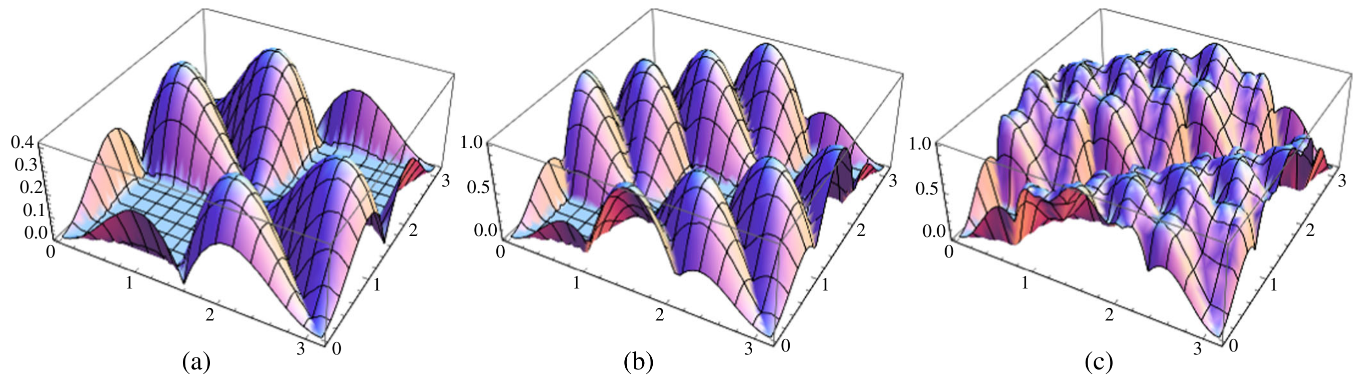

FIG. 1. Plots of the contextual fraction for empirical models obtained with projective measurements at ϕ_1 and ϕ_2 in the X-Y plane for each qubit on: (a) the Bell state $|\Phi^+\rangle = (1/\sqrt{2})(|00\rangle + |11\rangle)$; (b) $|\psi_{\mathsf{GHZ}(3)}\rangle$; (c) $|\psi_{\mathsf{GHZ}(4)}\rangle$, where the n-partite GHZ state ($n > 2$) is given by $|\Phi^+\rangle = (1/\sqrt{2})(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})$.

mixing with a weight $\lambda \in [0,1]$. The second is *controlled choice*: Given empirical models e and e' on scenarios $\langle X, \mathcal{M}, O\rangle$ and $\langle X', \mathcal{M}', O\rangle$, respectively, $e \,\&\, e'$ is defined on the scenario $\langle X \sqcup X', \mathcal{M} \sqcup \mathcal{M}', O\rangle$ by $(e \,\&\, e')_C \coloneqq e_C$ for $C \in \mathcal{M}$ and $(e \,\&\, e')_{C'} \coloneqq e_{C'}$ for $C' \in \mathcal{M}'$. The third is a *product*: $e \otimes e'$ is an empirical model defined on the scenario $\langle X \sqcup X', \mathcal{M} \star \mathcal{M}', O\rangle$, where $\mathcal{M} \star \mathcal{M}' \coloneqq \{C \sqcup C' | C \in \mathcal{M}, C' \in \mathcal{M}'\}$, by $(e \otimes e')_{C \sqcup C'}\langle s, s'\rangle \coloneqq e_C(s)e'_{C'}(s')$ for all $C \in M$, $C' \in \mathcal{M}'$, $s \in O^C$, and $s' \in O^{C'}$.

These operations can be used to construct any local empirical model on Bell scenarios starting from a very simple "generator": a deterministic model over a single measurement. This is illustrated in Supplemental Material [28].

Theorem 2. The contextual fraction is invariant under relabelings and nonincreasing under translation of measurements and coarse-graining of outcomes. For the combining operations, it satisfies the following properties: (i) $\mathsf{CF}(\lambda e + (1-\lambda)e') \leq \lambda \mathsf{CF}(e) + (1-\lambda)\mathsf{CF}(e)$, (ii) $\mathsf{CF}(e \,\&\, e') = \max\{\mathsf{CF}(e), \mathsf{CF}(e')\}$, (iii) $\mathsf{CF}(e \otimes e') = \mathsf{CF}(e) + \mathsf{CF}(e') - \mathsf{CF}(e)\mathsf{CF}(e')$.

A consequence of this result is that, for any of the combining operations, when e' is a noncontextual model (and thus composing with e' is a free operation), CF acts as a monotone: The contextual fraction of the new model is at most that of e (in fact, with equality holding for both choice and product).

Computational explorations.—General computational tools in the form of a *Mathematica* package have been developed implementing the two LPs described above to calculate, for any empirical model in any scenario: the (non) contextual fraction, a decomposition of the form (2), and the generalized Bell inequality from Theorem 1 (ii) for which the maximal violation is achieved. The package also calculates quantum empirical models from any (pure or mixed) state and any specified sets of compatible measurements.

As an example to illustrate the use of this package, we consider the empirical models obtained from local measurements on various n-qubit states. On each qubit, we allow the same two local measurements, equatorial on the Bloch sphere, parametrized by angles ϕ_1 and ϕ_2. Figure 1 plots the contextual fraction of the resulting models as a function of these angles.

Computational explorations of this kind can be a useful tool for guiding research, pointing the way to conjectures and results (e.g., [29–31]). A more detailed analysis of the examples from Fig. 1, leading to the characterization of a family of strongly contextual models arising from n-partite GHZ states, can be found in Supplemental Material [28].

Applications to quantum computation.—Contextuality has been associated with quantum advantage in certain information-processing and computational tasks. One use for a measure of contextuality is to quantify such advantages.

A computational model in which contextuality has been associated with an advantage is measurement-based quantum computation (MBQC). An $l2$-MBQC is a process with m classical bits of input and l of output, using an $(n,2,2)$ empirical model (n parties, two measurement settings per party, two outcomes per measurement) as a resource. The classical control—which preprocesses the inputs, determines the flow of measurements by choosing which sites to measure next and with which measurement setting (potentially depending on previous outcomes), and postprocesses to produce the outputs—can perform only $\mathbb{Z}_2$-linear computations. The additional power to compute nonlinear functions thus resides in certain resource empirical models.

In Ref. [2], Theorem 2, it was shown that if an $l2$-MBQC process deterministically calculates a non-$\mathbb{Z}_2$-linear Boolean function $f: 2^m \to 2^l$, then the resource is necessarily strongly contextual. A probabilistic version was also obtained in Ref. [2], Theorem 3: Contextuality must be present whenever a nonlinear function is calculated with a sufficiently large probability of success. By analyzing that proof, we extract a sharpened version of this result establishing a precise relationship between the hardness (nonlinearity) of the function, the probability of success, and the contextual fraction.

The *average distance* between two Boolean functions $f, g: 2^m \to 2^l$ is given by $\tilde{d}(f,g) \coloneqq 2^{-m}|\{\mathbf{i} \in 2^m | f(\mathbf{i}) \neq g(\mathbf{i})\}|$. The average distance of f to the closest $\mathbb{Z}_2$-linear function is denoted by $\tilde{\nu}(f)$.

Theorem 3. Let $f: 2^m \to 2^l$ be a Boolean function and consider an l2-MBQC that uses the empirical model e to compute f with average success probability $\bar{p}_S$ over all 2^m possible inputs and corresponding average failure probability $\bar{p}_F = 1 - \bar{p}_S$. Then, $\bar{p}_F \geq \mathsf{NCF}(e)\tilde{\nu}(f)$.

Note that for deterministic computation ($\bar{p}_S = 1$) of a nonlinear function ($\tilde{\nu} > 0$), we require strong contextuality [$\mathsf{NCF}(e) = 0$], recovering the deterministic result in Ref. [2]. More generally, for a given nonlinear function, the higher the desired success probability, the larger the contextual fraction must be. Additional details, including a rigorous presentation and proof, may be found in Supplemental Material [28].

Similar results can be obtained to quantify advantage in games, generalizing non-local games on Bell scenarios [32–36]. A game is specified by n Boolean formulas, one for each context, which describe the winning condition that the output must satisfy. If the formulas are k-consistent, meaning that at most k of them have a joint satisfying assignment, then the hardness of the game is measured by $(n-k)/n$. One can show that $\bar{p}_F \geq \mathsf{NCF}(e)(n-k)/n$, relating the probability of success, the noncontextual fraction, and the hardness of the task. See [37] for the relation with Bell inequalities, from which a proof of this result follows. Details are given in Supplemental Material, Theorem 4. Further development of these ideas is a topic for future research.

This work was done in part while the authors visited the Simons Institute for the Theory of Computing (supported by the Simons Foundation) at the University of California, Berkeley, as participants of the Logical Structures in Computation program, and while S. M. was based at l'Institut de Recherche en Informatique Fondamentale, Université Paris Diderot (Paris 7). Support from the following is also gratefully acknowledged: EPSRC EP/N018745/1, Contextuality as a Resource in Quantum Computation (S. A. and R. S. B.); Fondation Sciences Mathématiques de Paris, postdoctoral research grant eotpFIELD15RPOMT-FSMP1, Contextual Semantics for Quantum Theory (S. M.); the Oxford Martin School James Martin Program on Bio-inspired Quantum Technologies (S. A., S. M., and R. S. B.); FCT, Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology), Ph.D. Grant No. SFRH/BD/94945/2013 (R. S. B.).

[1] J. Anders and D. E. Browne, Computational Power of Correlations, Phys. Rev. Lett. **102**, 050502 (2009).
[2] R. Raussendorf, Contextuality in measurement-based quantum computation, Phys. Rev. A **88**, 022322 (2013).
[3] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Contextuality supplies the 'magic' for quantum computation, Nature (London) **510**, 351 (2014).
[4] S. Abramsky and A. Brandenburger, The sheaf-theoretic structure of non-locality and contextuality, New J. Phys. **13**, 113036 (2011).
[5] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).
[6] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Nonlocal correlations as an information-theoretic resource, Phys. Rev. A **71**, 022101 (2005).
[7] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Operational Framework for Nonlocality, Phys. Rev. Lett. **109**, 070401 (2012).
[8] J. I. de Vicente, On nonlocality as a resource theory and nonlocality measures, J. Phys. A **47**, 424017 (2014).
[9] See also the recent Ref. [10], which develops a different approach to a resource theory for contextuality.
[10] B. Amaral, A. Cabello, M. Terra Cunha, and L. Aolita, Noncontextual wirings, arXiv:1705.07911.
[11] These include a negative probability measure introduced in Ref. [4], contextuality-by-default measures [12], and various other measures considered for nonlocality in specific measurement scenarios, e.g., noise-based [13–15] and inefficiency-based [16–18] measures, known to relate to communication complexity [18,19].
[12] E. N. Dzhafarov, J. V. Kujala, and V. H. Cervantes, in *Proceedings of the International Symposium on Quantum Interaction* (Springer, New York, 2015), pp. 12–23.
[13] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski, and A. Zeilinger, Violations of Local Realism by Two Entangled N-Dimensional Systems Are Stronger than for Two Qubits, Phys. Rev. Lett. **85**, 4418 (2000).
[14] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, Quantum nonlocality in two three-level systems, Phys. Rev. A **65**, 052325 (2002).
[15] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. M. Wolf, Operator Space Theory: A Natural Framework for Bell Inequalities, Phys. Rev. Lett. **104**, 170405 (2010).
[16] S. Massar, S. Pironio, J. Roland, and B. Gisin, Bell inequalities resistant to detector inefficiency, Phys. Rev. A **66**, 052112 (2002).
[17] S. Laplante, V. Lerays, and J. Roland, in *Automata, Languages, and Programming: 39th International Colloquium (ICALP 2012), Proceedings, Part I*, edited by A. Czumaj, K. Mehlhorn, A. Pitts, and R. Wattenhofer, Lecture Notes in Computer Science Vol. 7391 (Springer, New York, 2012), pp. 617–628.
[18] J. Degorre, M. Kaplan, S. Laplante, and J. Roland, in *Proceedings of the 34th International Symposium on Mathematical Foundations of Computer Science (MFCS 2009)*, edited by R. Královič and D. Niwiński, Lecture Notes in Computer Science Vol. 5734 (Springer, 2009), pp. 270–281.
[19] S. Laplante, M. Laurière, A. Nolin, J. Roland, and G. Senno, in *Proceedings of the 11th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2016)*, edited by A. Broadbent, Leibniz International

Proceedings in Informatics (LIPIcs) Vol. 61 (Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016), pp. 5:1–5:24.
[20] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).
[21] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).
[22] A subprobability distribution on a set S is a map $b: S \to \mathbb{R}_{\geq 0}$ with finite support and $w(b) \leq 1$, where $w(b) \coloneqq \sum_{s \in S} b(s)$ is called its weight. The set of subprobability distributions on S is ordered pointwise: b' is a subdistribution of b (written $b' \leq b$) whenever $\forall s \in S, b'(s) \leq b(s)$.
[23] A. C. Elitzur, S. Popescu, and D. Rohrlich, Quantum nonlocality for each pair in an ensemble, Phys. Lett. A **162**, 25 (1992).
[24] See also [25,26], where the term local fraction is actually used.
[25] J. Barrett, A. Kent, and S. Pironio, Maximally Nonlocal and Monogamous Quantum Correlations, Phys. Rev. Lett. **97**, 170409 (2006).
[26] L. Aolita, R. Gallego, A. Acín, A. Chiuri, G. Vallone, P. Mataloni, and A. Cabello, Fully nonlocal quantum correlations, Phys. Rev. A **85**, 032107 (2012).
[27] G. B. Dantzig and M. N. Thapa, *Linear programming 2: Theory and Extensions*, Springer Series in Operations Research and Financial Engineering (Springer Verlag, Berlin, 2003).
[28] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.119.050504 for rigorous proofs and further details on our computational explorations.
[29] S. Abramsky, C. M. Constantin, and S. Ying, Hardy is (almost) everywhere: Nonlocality without inequalities for almost all entangled multipartite states, Inf. Comput. **250**, 3 (2016).
[30] S. Abramsky and C. Constantin, A classification of multipartite states by degree of non-locality, Electron. Proc. Theor. Comput. Sci. **171**, 10 (2014).
[31] S. Mansfield, D.Phil. thesis, University of Oxford, 2013.
[32] R. Cleve, P. Høyer, B. Toner, and J. Watrous, in *Proceedings of the 19th IEEE Annual Conference on Computational Complexity (CCC 2004)* (IEEE, Piscataway, New Jersey, 2004), pp. 236–249.
[33] R. Cleve and R. Mittal, in *Automata, Languages, and Programming: 41st International Colloquium (ICALP 2014), Proceedings, Part I*, edited by J. Esparza, P. Fraigniaud, T. Husfeldt, and E. Koutsoupias, Lecture Notes in Computer Science Vol. 8572 (Springer, New York, 2014), pp. 320–331.
[34] L. Mančinska and D. E. Roberson, Quantum homomorphisms, J. Comb. Theory B **118**, 228 (2016).
[35] R. Cleve, L. Liu, and W. Slofstra, Perfect commuting-operator strategies for linear system games, J. Math. Phys. (N.Y.) **58**, 012202 (2017).
[36] S. Abramsky, R. S. Barbosa, N. de Silva, and O. Zapata, *Proceedings of the 42nd International Symposium on Mathematical Foundations of Computer Science (MFCS 2017)* (2017) [arXiv:1705.07310].
[37] S. Abramsky and L. Hardy, Logical Bell inequalities, Phys. Rev. A **85**, 062114 (2012).