Contextuality in measurement-based quantum computation

Robert Raussendorf*

Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

(Received 1 May 2013; revised manuscript received 11 July 2013; published 19 August 2013)

We show, under natural assumptions for qubit systems, that measurement-based quantum computations (MBQCs) which compute a nonlinear Boolean function with a high probability are contextual. The class of contextual MBQCs includes an example which is of practical interest and has a superpolynomial speedup over the best-known classical algorithm, namely, the quantum algorithm that solves the "discrete log" problem.

DOI: 10.1103/PhysRevA.88.022322 PACS number(s): 03.67.Ac, 03.65.Ta

I. INTRODUCTION

While numerous quantum algorithms have been found that offer polynomial or superpolynomial speedups over their classical counterparts [1–3], the precise quantum mechanical origin of this speedup remains unknown. The prominent candidates—entanglement [4], superposition and interference [5], and largeness of Hilbert space—provide an intuitive understanding in many situations. Yet, as a whole, the phenomenology so far uncovered does not lend itself to a simple interpretation [6–12].

Here we turn our attention to a different characterization of nonclassicality, namely, contextuality [13,14], and study its relation to computational power. We choose measurement-based quantum computation (MBQC) [15] as our setting. The starting point for this investigation is the observation by Anders and Browne [16] that one of Mermin's proofs [17] of the Kochen-Specker theorem [13] can be converted into a simple MBQC. We are led to ask whether the connection between MBQC and contextuality exhibited by this example is accidental or whether it holds in general. The main finding of this paper is that, under quite natural assumptions for multiqubit systems, all MBQCs which compute a nonlinear Boolean function with a sufficiently high success probability are contextual.

For MBQC, the separation between linear and nonlinear functions is fundamental. Every MBQC requires a classical control computer for adjusting measurement bases according to the computational input and for converting measurement outcomes into computational output. This classical side processing is all linear. Evaluating nonlinear functions is out of reach for such a classical control computer without access to additional resources.

This paper is organized as follows. In Sec. II, we review Anders and Browne's example and define the setting of MBQC and notions of contextuality we use. In Sec. III we present three results on the interplay between contextuality and the nonlinearity of the computational output, Theorems 2, 3, and 5 . We point out that the class of contextual MBQCs contains a computation which is of actual algorithmic interest, i.e., achieves a superpolynomial speedup over the best-known classical algorithm. It is the MBQC variant of the quantum algorithm for the "discrete log" problem [1,18]. In Sec. IV, we discuss experimental tests of contextuality. We conclude with a discussion in Sec. V.

II. THE SETTING

We discuss the link between contextuality and quantum computation for MBQC [15]. MBQC is a model of quantum computation in which a quantum algorithm is implemented solely by local measurements on a fixed initial state. The choice of measurement bases determines the algorithm to be implemented, and correlations among the measurement outcomes reveal the result of the computation. The computational power of this scheme is fully determined by the initial quantum state.[1] For suitable initial states such as cluster states, MBQC is universal.

A. Computation and contextuality: A first example

Following Anders and Browne [16], we consider a three-party Greenberger-Horne-Zeilinger (GHZ) [21] state $|\mathrm{GHZ}\rangle = \frac{|000\rangle+|111\rangle}{\sqrt{2}}$, which can be used to execute a deterministic OR gate within the framework of MBQC. While standard electronic devices routinely perform OR gates without quantum-mechanical action, this result offers a structural insight into MBQC. Namely, it is known that every MBQC requires a classical control computer that converts the classical input into measurement settings and the measurement outcomes into computational output. This classical control computer is capable of doing only one type of operation: addition mod 2. It is thus not classically universal and, indeed, very limited. Now, having access to GHZ states and local projective measurements promotes this control computer to classical universality. Thus, in the described setting, the access to quantum resources vastly increases the set of computable functions.

What is more, Anders and Browne's construction repurposes an existing proof [17] of the Kochen-Specker theorem [13] into a quantum mechanical computation. The computation takes two bits of input, i_1 and i_2, and outputs a single bit $o \equiv i_1 \vee i_2$. It proceeds as follows. *Step 1*: The settings for the local measurements on the three qubits are calculated from the input i_1 and i_2. For either of the three qubits, *a priori* the Pauli

*rraussendorf@phas.ubc.ca

[1]Note, however, that other schemes of universal quantum computation by measurement exist in which the measurements are *not* local [19,20]. For such schemes, the initial quantum state of the system is irrelevant.

observables $O_k = X_k, Y_k$ can be measured (here and in the following, $\sigma_x \equiv X$, $\sigma_y \equiv Y$, $\sigma_z \equiv Z$), and we use the binary variable q_k to encode the choice. If $q_k = 0$ (1), then X_k (Y_k) is measured. The measurement setting $\mathbf{q} = (q_1, q_2, q_3)$ is related to the input $\mathbf{i} = (i_1, i_2)$ via $q_1 = i_1, q_2 = i_2, q_3 = i_1 + i_2 \bmod 2$. *Step 2*: The observables $O_k(q_k)$ are being measured, whereby the measurement outcomes $s_k \in \{0,1\}$ are obtained. Here, if the observed value of the Pauli observable O_k was $+1$ (-1), then $s_k = 0$ ($s_k = 1$). *Step 3*: The parity $o \equiv s_1 + s_2 + s_3 \bmod 2$ of the three measurement outcomes is computed and outputted.

It is easily verified that this procedure does indeed compute the desired OR gate. First note that $|\mathrm{GHZ}\rangle$ is an eigenstate with eigenvalue 1 of the following operators:

$$X_1X_2X_3, \quad -X_1Y_2Y_3, \quad -Y_1X_2Y_3, \quad -Y_1Y_2X_3. \tag{1}$$

The outcomes of the local X and Y measurements selected by the input $\mathbf{i}$ in the above procedure are thus strictly correlated or anticorrelated. Specifically, if $i_1 = i_2 = 0$, then the measured observables are X_1, X_2, X_3. The measurement outcomes s_1, s_2, s_3 are individually random, but because of $X_1X_2X_3|\mathrm{GHZ}\rangle = |\mathrm{GHZ}\rangle$, $o(0,0) = s_1 + s_2 + s_3 \bmod 2 = 0$ with certainty. Likewise, if $i_1 = 0$, $i_2 = 1$, then the measured observables are X_1, Y_2, Y_3. As before, the local measurement outcomes s_1, s_2, s_3 are individually random, but because of the relation $X_1Y_2Y_3|\mathrm{GHZ}\rangle = -|\mathrm{GHZ}\rangle$ we find that $o(0,1) = s_1 + s_2 + s_3 \bmod 2 = 1$ with certainty. The remaining two cases are analogous, and we thus verify the logical table of the OR gate.

The present implementation of the OR gate is quantum mechanical, and one may ask whether contextuality of quantum mechanics is brought to bear in this process. Let us try to construct a noncontextual hidden variable model (HVM) for the "predetermined" measurement outcomes x_1, x_2, x_3, y_1 y_2, $y_3 \in \mathbb{Z}_2$ of the observables X_1, X_2, X_3, Y_1, Y_2, Y_3 potentially measured in the realization of the gate. Since the measured observables are all local, such an HVM is also local (a special case of being noncontextual [17]). Going through each entry in the logical table of the OR gate, the following relations are imposed:

$$\begin{aligned} x_1 + x_2 + x_3 \quad &\bmod 2 = 0, \\ x_1 + y_2 + y_3 \quad &\bmod 2 = 1, \\ y_1 + x_2 + y_3 \quad &\bmod 2 = 1, \\ y_2 + y_2 + y_3 \quad &\bmod 2 = 1. \end{aligned} \tag{2}$$

Adding up these four equations, we find $0 = 2(x_1 + y_1 + x_2 + y_2 + x_3 + y_3) \bmod 2 = 1$. Contradiction! No assignment of predetermined local measurement outcomes reproduces the correlations required for an OR gate in the present three-party setting. The argument we have just stated is, in fact, Mermin's state-dependent proof of the Kochen-Specker theorem in dimension 8 [17]. We find that a proof of contextuality of quantum mechanics can be repurposed as a (simple) quantum computation.

We may take this example a step further and consider the following modifications: (i) flipping some observables $O_k(q_k) \longrightarrow -O_k(q_k)$ and (ii) using, instead of $|\mathrm{GHZ}\rangle$, some other state from the GHZ family, i.e., a simultaneous eigenstate of the observables in Eq. (1), but with eigenvalues -1 for some of them. Since both the changes can be implemented by local unitary operation, one expects contextuality to remain unaffected. And indeed, while the right-hand side of Eq. (2) does change under such transformations, the number of entries 1 always remains odd. Hence, the contradiction for noncontextual HVMs persists. As for the computed Boolean function, it also changes but always remains nonlinear. We are thus led to ask: Is there a link between the nonlinearity of a Boolean function computed in MBQC and the noncontextuality of such a computation? This is the question which we subsequently investigate. To do so, we first need to define our precise setting of MBQC and notion of contextuality.

B. The general setting of measurement-based computation

The above example using a GHZ state may serve as a first illustration of MBQC, but it misses two aspects: (i) MBQC is universal for quantum computation, and (ii) the measurements in MBQCs can be temporally ordered. The latter is a consequence of the randomness inherent in quantum measurement. To prevent this randomness from creeping into the logical processing, measurement bases need to be adjusted to measurement outcomes already obtained. This leads to a partial temporal order of the measurement events; see Fig. 2 in the Appendix.

One may consider an MBQC scenario with n parties, k measurement settings at each party, and l possible outcomes for each of those measurements. However, for the confines of this paper we restrict our attention to the case of two measurement settings per party and two outcomes for each local measurement, i.e., $k = l = 2$. This is a natural choice when the local quantum systems are qubits. We further impose a restriction on the classical side processing in MBQC. Side processing is required for adjusting the measurement bases according to previously obtained outcomes and for obtaining the computational output from the local measurement outcomes. We require all such processing to be addition mod 2.

The relations between contextuality and computational power described in Sec. III (or at least their present proofs) crucially depend on this linearity. Therefore, before making definitions, we need to motivate such linear constraints. In this regard, we note that in the GHZ example in the previous section all classical side processing is indeed mod 2 linear. However, the main justification for imposing mod 2 linear relations of classical side processing is that they are sufficient for quantum mechanically universal MBQC on cluster states [15]. The origin of linearity in the classical side processing is explained in the Appendix. We note that MBQC schemes with different classical processing exist [22].

We now introduce the notion of *l2-MBQC* for "MBQC with mod 2 linear classical processing".

Definition 1. A l2-MBQC is an MBQC with classical input $\mathbf{i}$ and classical output $\mathbf{o}$, where the measurements driving the computation are all local and satisfy the following properties:

(1) For each party k, $k = 1 \ldots n$, there is a binary choice for the measurement basis, $q_k \in \{0,1\}$.

(2) For each party k and each $q_k \in \{0,1\}$, the measurement outcome is binary valued, $s_k \in \{0,1\}$.

(3) The computational output $\mathbf{o}$ is bitwise a parity of measurement outcomes $\mathbf{s} = (s_1, s_2, \ldots, s_n)^T$,

$$\mathbf{o} = Z\mathbf{s} \quad \bmod 2. \tag{3}$$

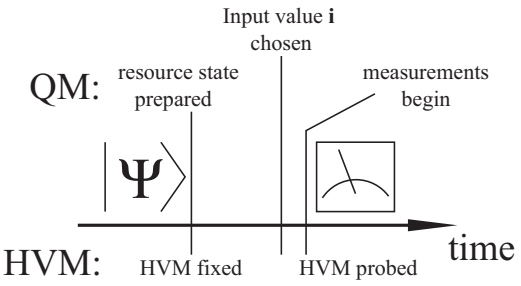

FIG. 1. Timeline of events in an MBQC. (1) The resource quantum state is prepared. (2) The computational input **i** is chosen. (3) The measurements comprising the computation are performed, and their outcomes are processed. The hidden variable model attempting to describe the computation is fixed *before* the input **i** is chosen.

(4) The choice of measurement bases $\mathbf{q} = (q_1,q_2,q_n)^T$ is related to the measurement outcomes **s** and the binary-valued classical input $\mathbf{i} = (i_1,\ldots,i_l)^T$ via

$$\mathbf{q} = T\mathbf{s} + Q\mathbf{i} \mod 2. \tag{4}$$

(5) For a suitable ordering of the parties $1\ldots n$, the matrix T in Eq. (4) is lower triangular with vanishing diagonal.

The reason for imposing condition 5 is that if and only if T is strictly lower triangular with respect to a suitable ordering of parties, the given MBQC is *runnable*, i.e., measurement bases depend only on measurement outcomes that have already been obtained.

C. Contextuality

In an HVM, in stark contrast to quantum mechanics, measurement outcomes exist prior to measurement and are merely "revealed." Noncontextuality means the following: Let an observable A be measured jointly with one of the compatible observables B or C, and B be incompatible with C. An HVM is noncontextual if the "pre-existing" measurement outcome $\lambda(A)$ for A is independent of whether A is measured jointly with B or with C. Noncontextual HVMs cannot reproduce all predictions of quantum mechanics in Hilbert spaces of dimension $\geqslant 3$. This is the content of the Kochen-Specker theorem [13].

Noncontextuality is not compromised by the classical communication required in l2-MBQC. In quantum mechanics, two observables are compatible if and only if the corresponding Hermitian operators commute. Consider two parties, a and b, with to-be-measured observables, $O_a(q_a)$ and $O_b(q_b)$, where $q_a,q_b \in \{0,1\}$. The values of q_a and q_b are specified by prior measurement outcomes; see Eq. (4) and Fig. 1. Independent of the values of q_a and q_b, the observables O_a and O_b always commute because they are local to different tensor product factors of the underlying Hilbert space.

We follow the sheaf-theoretic notion of contextuality developed by Abramsky and Brandenburger [23]. Below we restate from [23] the definitions of the notions required for the present discussion, namely, "section," "measurement context," 'phenomenological model," "strongly contextual," and "contextual." (Note: In [23], "contextual" is called "probabilistically nonextendable.") We specialize to the case where all measurement outcomes are binary.

Sections. Let X be the set of measurements and $\mathbb{Z}_2$ the set of outcomes for each individual measurement. For all $U \subseteq X$, a section over U is a function $s : U \to \mathbb{Z}_2$. It describes measurement outcomes $s = (s(O_1),s(O_2),\ldots,s(O_n))$, $O_i \in U$. We denote by $\mathcal{E}(U)$ the set of sections s over U. A section over X is called "global." Contextuality is about the nonexistence of global sections.

Measurement contexts. A measurement context is a set $C \subset X$ of compatible measurements. The set $\mathcal{M}$ of measurement contexts has the following two properties: (i) $X = \bigcup_{C\in\mathcal{M}} C$, and (ii) for $C,C' \in \mathcal{M}$, if $C \subseteq C'$, then $C = C'$. The second property says that contexts are maximal sets of compatible measurements.

A phenomenological model e is a set of probability distributions $\{e_C | C \in \mathcal{M}\}$ such that for each measurement context C and measurement outcome $\mathbf{s} \in \mathcal{E}(C)$, $e_C(\mathbf{s})$ is the probability of obtaining **s** within C. We may consider any process that begins with a preparation and ends with a measurement—quantum, classical, or other—as a phenomenological model.

Strong contextuality. We define the set $\mathcal{S}_e$ of global sections that only predict possible events:

$$\mathcal{S}_e := \{\mathbf{s} \in \mathcal{E}(X) | \, e_C(\mathbf{s}|_C) > 0, \forall C \in \mathcal{M}\}. \tag{5}$$

Definition 2. A phenomenological model e is strongly contextual if $\mathcal{S}_e = \emptyset$.

An example of a strongly contextual model is the GHZ scenario discussed in Sec. II A.

A model e with $\mathcal{S}_e = \emptyset$ is definitely contextual since no assignment $\mathbf{s} \in \mathcal{E}(X)$ of "predetermined" measurement outcomes—and no probability distribution over such assignments—can reproduce it. Such contextuality is called "strong" because it implies other forms of contextuality [23].

Contextuality. A phenomenological model e may be contextual even if $\mathcal{S}_e \neq \emptyset$. While consistent assignments for the predetermined measurement outcomes exist, no probability distribution over those assignments may reproduce the probability distributions in e.

We label the elements in $\mathcal{S}_e$ by a hidden variable λ, i.e., $\mathcal{S}_e = \{\mathbf{s}(\lambda),\lambda \in \Lambda\}$. Each $\mathbf{s}(\lambda)$ induces a set of probability distributions $\{d_C(\lambda),C \in \mathcal{M}\}$ over the measurement records in all measurement contexts. Then a probability distribution $\tilde{p}$ of the hidden variable λ induces a set of probability distributions $\{d_C(\tilde{p}) = \sum_{\lambda\in\Lambda} \tilde{p}(\lambda)\, d_C(\lambda),C \in \mathcal{M}\}$ over the measurement records in all measurement contexts.

Definition 3. A phenomenological model e is contextual if the set of linear equations,

$$e_C = \sum_{\lambda\in\Lambda} \tilde{p}(\lambda)\, d_C(\lambda), \quad \forall C \in \mathcal{M},$$

has no solution $\tilde{p}$ with $\tilde{p}(\lambda) \geqslant 0$, for all $\lambda \in \Lambda$, and subject to the normalization constraint $\sum_\lambda \tilde{p}(\lambda) = 1$. If it has such a solution, the model is noncontextual.

Contextuality is weaker than strong contextuality: the latter always implies the former. The converse does not hold. The Bell scenario is contextual but not strongly contextual [23].

Now consider a phenomenological model e with a noncontextual part K and a general no-signaling part Q:

$$e = p\,K + (1-p)Q, \quad 0 \leqslant p \leqslant 1. \tag{6}$$

That is, for all contexts $C \in \mathcal{M}$ and all sections $\mathbf{s} \in \mathcal{E}(C)$ we have $e_C(\mathbf{s}) = p\,K_C(\mathbf{s}) + (1-p)Q_C(\mathbf{s})$. We call the supremum

of p over all convex decompositions, Eq. (6), the noncontextual fraction of e. A phenomenological model e is called *maximally contextual* if its noncontextual fraction is 0. We then have the following relation [23].

Theorem 1. A model is strongly contextual if and only if it is maximally contextual.

This concludes our review of the required notions of contextuality from [23].

A feature of measurement contexts in l2-MBQC that is not explicitly addressed in [23] but can be included is the adaptivity of local measurement settings according to previously obtained measurement outcomes. In l2-MBQC, the measurement at party k depends on the m-bit input $\mathbf{i}$ and the measurement outcomes $s|_{\mathcal{P}(k)}$ obtained in the past $\mathcal{P}(k)$ of k; cf. Eq. (4).

Contexts are labeled by the basis choice $\mathbf{q}$:

$$C(\mathbf{q}) = \{O_k(q_k), k = 1, \ldots, n\}. \tag{7}$$

The measurement record $\mathbf{s}$ appearing in Eqs. (3) and (4) and the section s are related via

$$\mathbf{s} = s|_{C(\mathbf{q})}\,. \tag{8}$$

Since $\mathbf{q}$ depends on $\mathbf{s}$ via Eq. (4), this looks like a self-consistency condition. Given s and $\mathbf{q}$, it is *a priori* not clear whether $\mathbf{s}$ exists and, if it it does, whether it is unique. However, due to the runnability condition on matrix T in Eq. (4), a unique solution $\mathbf{s}$ does indeed always exist. The set $\{1, \ldots, n\}$ of qubits is divided into smaller sets $\mathcal{Q}_0$, $\mathcal{Q}_1, \ldots, \mathcal{Q}_{\text{last}}$, which are measured one after the other. The measurement bases for any given set only require knowledge of the measurement outcomes from prior sets. Choosing the bases for the measurements in $\mathcal{Q}_0$ requires no knowledge of measurement outcomes, and the set $\mathcal{Q}_0$ is therefore measured first. Using the measurement outcomes from $\mathcal{Q}_0$, $\mathbf{q}|_{\mathcal{Q}_1}$ can be obtained and $\mathcal{Q}_1$ be measured, and so on. Correspondingly, the components of $\mathbf{s}$ can be extracted one set $\mathcal{Q}_i$ at a time.

Remark 1. MBQC [15] uses only local observables to drive the computation. An alternative approach therefore is to relate MBQC to the nonlocality of quantum mechanics rather than contextuality. If so, a complication is posed by the adaptive choice of measurement bases in MBQC. The local parties exchange classical messages in order to adjust measurement, which runs counter to the assumption of locality. In this situation, postselection may be employed to restore locality. Furthermore, for the kind of postselection required, CHSH-type inequalities indicating the nonlocality of MBQCs can still be established [24]. However, the fraction of measurement records that survive the postselection criterion is, for MBQCs with temporal order, in general, exponentially small in the number n of parties.

III. CONTEXTUALITY AND COMPUTATIONAL POWER

In this section we present our results on the interplay between contextuality and nonlinearity in MBQC.

A. The deterministic case

We call a l2-MBQC *deterministic* if it computes the vector-valued Boolean function $\mathbf{o}$ with unit probability for every allowed input $\mathbf{i} \in \mathbb{Z}_2^m$. We regard l2-MBQCs as phenomenological models. That is, a l2-MBQC M is a set of probability distributions, $M = \{M_C | \, C \in \mathcal{M}\}$.

Theorem 2. Let M be a l2-MBQC which deterministically evaluates a vector $\mathbf{o}$ of Boolean functions on an input $\mathbf{i}$. If $\mathbf{o}(\mathbf{i})$ is nonlinear mod 2 in $\mathbf{i}$, then M is strongly contextual.

Proof of Theorem 2. We show that if M is not strongly contextual, then $\mathbf{o}(\mathbf{i})$ is mod 2 linear in $\mathbf{i}$. The result then follows by negation.

If M is not strongly contextual, then there exists a noncontextual HVM consistently assigning values to the observables in the set $X = \bigcup_{\mathbf{q}} C(\mathbf{q})$, i.e., $\mathcal{S}_M$ is nonempty. For each party $k = 1 \ldots n$, there are at most two possible measurement bases, labeled $q_k = 0$ and $q_k = 1$, respectively (property 1 in Definition 1). Therefore, $X \subset \tilde{X} := \{O_k(q_k = 0), O_k(q_k = 1), \forall k = 1 \ldots n\}$, but X may be strictly smaller than $\tilde{X}$. First, consider the case where both $O_k(q_k = 0)$ and $O_k(q_k = 1)$ are in X. In the noncontextual HVM, these observables have pre-existing outcomes $s_i(q_k = 0)$ and $s_i(q_k = 1)$ that are independent of the context $C(\mathbf{q})$. Since the measurement outcomes s_k are binary (property 2), and any function defined on only two points is linear, these outcomes can be expressed in terms of two binary variables, c_k and d_k:

$$s_k(q_k) \equiv c_k \oplus d_k q_k. \tag{9}$$

The pair $[s_k(q_k = 0), s_k(q_k = 1)]$ and the pair $[c_k, d_k]$ contain the exact same information. Keep in mind that c_k and d_k depend upon the chosen $s \in \mathcal{S}_M$.

Next, consider the case where, for a given party k, only one of the two values of q_k occurs for all $\mathbf{s} \in \mathcal{S}_M$ and all $\mathbf{i} \in \mathbb{Z}_2^m$. This can happen only if the kth row of Q is identically 0. Then Eq. (9) does still hold. The only difference is that c_k, d_k are no longer unique.

Thus, the relation Eq. (9) holds for all parties $k = 1 \ldots n$. We convert it into vector form, $\mathbf{s}(\mathbf{q}) \equiv \mathbf{c} \oplus D\mathbf{q}$, where $D = \text{diag}(d_k)$. Inserting Eq. (4) (see property 4 of Definition 1) into this relation, we obtain

$$(I \oplus D\,T)\mathbf{s} \equiv \mathbf{c} \oplus D\,Q\,\mathbf{i}.$$

Therein, the matrix $I \oplus D\,T$ is invertible because of property 5. We can thus always solve for $\mathbf{s}$, and using Eq. (3) (see property 3 in Definition 1), we obtain for the classical output $\mathbf{o}$ of the computation

$$\mathbf{o} = \mathbf{c}' + Q'\,\mathbf{i} \mod 2, \tag{10}$$

where $\mathbf{c}' = Z(I \oplus DT)^{-1}\mathbf{c}$ and $Q' = Z(I \oplus DT)^{-1}DQ$. We emphasize that $\mathbf{c}'$ and Q' may depend on the choice $s \in \mathcal{S}_M$ via $\mathbf{c}$ and D; cf. Eq. (9). Therefore, $\mathbf{o}(\mathbf{i})$ is linear in $\mathbf{i}$ given a particular $s \in \mathcal{S}_M$. To make explicit the potential dependence of $\mathbf{o}$ on the choice of the global section, we choose a reference section $s_0 \in \mathcal{S}_M$; hence $\mathbf{o}(s_0) = \mathbf{c}'(s_0) + Q'(s_0)\,\mathbf{i} \bmod 2$. Now, for any choice $s_\lambda \in \mathcal{S}_M$, for all $\mathbf{i} \in \mathbb{Z}_2^m$ we must have

$$\mathbf{o}(s_0) = \mathbf{c}'(s_0) \oplus Q'(s_0)\mathbf{i} = \mathbf{c}'(s_\lambda) \oplus Q'(s_\lambda)\mathbf{i} = \mathbf{o}(s_\lambda);$$

otherwise, the computation would not be deterministic. Thus, for all $s_\lambda \in \mathcal{S}_M$, $\mathbf{c}'(s_\lambda) = \mathbf{c}'(s_0) =: \mathbf{c}'$ and $Q'(s_\lambda) = Q'(s_0) =: Q'$. The output $\mathbf{o}(\mathbf{i})$ in Eq. (10) is thus linear mod 2 in $\mathbf{i}$. ■

It should be noted that, besides the simple initial example of [16] (cf. Sec. II A), the class of contextual l2-MBQCs contains a quantum algorithm with superpolynomial speedup over the best-known classical counterpart. Namely, the quantum

algorithm solving the "discrete log" problem can be made deterministic in the circuit model [18], and its translation as a l2-MBQC thus falls under Theorem 2.

B. The probabilistic case

Since quantum-mechanical phenomena are in general statistical rather than deterministic, one may be interested in probabilistic extensions of Theorem 2. To begin, we need a notion of probabilistic function evaluation.

Definition 4. A procedure τ probabilistically evaluates a vector-valued Boolean function $\mathbf{o}(\mathbf{i})$ on an m-bit input $\mathbf{i} \in \mathbb{Z}_2^m$ with success probability p_S if

$$\min_{\mathbf{i}\in\mathbb{Z}_2^m} \mathrm{Prob}[\tau(\mathbf{i}) = \mathbf{o}(\mathbf{i})] = p_S.$$

The realization of probabilistic function evaluation as a l2-MBQC has the phenomenological model

$$M = pK + (1-p)Q, \tag{11}$$

where Q is a contextual no-signaling model and K a noncontextual model, such that the noncontextual fraction p is maximized (subject to the constraint $0 \leqslant p \leqslant 1$). The model M is contextual for all $p < 1$ and, with Theorem 1, strongly contextual for $p = 0$. Now, consider the case where M deterministically evaluates a nonlinear Boolean function, $p_S = 1$. By Theorem 2, M is then strongly contextual. Hence, by Theorem 1, M is maximally contextual, $p = 0$. We now ask, For probabilistic evaluation of a Boolean function with a l2-MBQC M, how much can the success probability p_S drop for the computation to remain contextual ($p < 1$)? This question leads us to the following theorem.

Theorem 3. Let M_{p_s} be a l2-MBQC that probabilistically evaluates a vector of nonlinear Boolean functions on m bits of input, with success probability p_S. If $p_S > 1 - \frac{1}{2^m}$, then M_{p_S} is contextual.

We thus find that contextuality persists within a finite interval around the point of strong contextuality. However, we also find that for general nonlinear functions, the contextuality threshold for p_S approaches unity exponentially fast in the input size m. This result can be significantly improved for special nonlinear Boolean functions, as discussed below.

In preparation for the proof of Theorem 3, we define the distance ν of a Boolean function o to the closest linear Boolean function, $\nu = \min_{l\in\text{lin.B.f}} \mathrm{wt}(o \oplus l)$. For a vector $\mathbf{o}$ of Boolean functions, we define the distance to the closest linear function as

$$\nu := \min_{\mathbf{l}\in\text{lin.B.f}} \sum_{\mathbf{i}\in\mathbb{Z}_2^m} 1 - \delta(\mathbf{o}(\mathbf{i}) \oplus \mathbf{l}(\mathbf{i})).$$

We then have the following.

Lemma 1. Let M_{p_S} be a l2-MBQC that evaluates with success probability p_S a vector-valued Boolean function on m input bits with distance ν to the closest linear function. If $p_S > 1 - \frac{\nu}{2^m}$, then M_{p_S} is contextual.

Proof of Lemma 1. We decompose the noncontextual part pK on the right-hand side of Eq. (11) as

$$pK = \sum_k p_k L_k,$$

where all L_k are noncontextual models corresponding to l2-MBQCs which deterministically evaluate functions $\mathbf{l}_k$, and all $p_k \geqslant 0$. By Theorem 2, all functions $\mathbf{l}_k$ are linear mod 2. We define $\mathcal{L}_k := \{\mathbf{i} \in \mathbb{Z}_2^m \,|\, l_k(\mathbf{i}) \neq \mathbf{o}(\mathbf{i})\}$. Then

$$\nu \leqslant |\mathcal{L}_k|, \quad \forall k. \tag{12}$$

For any given input $\mathbf{i} \in \mathbb{Z}_2^m$, the noncontextual part pK in Eq. (11) contributes a portion $p_{\text{fail},L}(\mathbf{i})$ to the probability of failure to output $\mathbf{o}(\mathbf{i})$, $p_{\text{fail},L}(\mathbf{i}) = \sum_{k|\mathbf{i}\in\mathcal{L}_k} p_k$. The contextual part $(1-p)Q$ in Eq. (11) may also contribute, and the probability $p_{\text{fail}}(\mathbf{i})$ of failure to output $\mathbf{o}(\mathbf{i})$ is thus the same or bigger. Summing over all $\mathbf{i} \in \mathbb{Z}_2^m$, we have

$$\sum_{\mathbf{i}} p_{\text{fail}}(\mathbf{i}) \geqslant \sum_{\mathbf{i}} \sum_{k|\mathbf{i}\in\mathcal{L}_k} p_k = \sum_k p_k |\mathcal{L}_k| \geqslant p\nu. \tag{13}$$

Further, $1 - p_S = \max_{\mathbf{i}\in\mathbb{Z}_2^m} p_{\text{fail}}(\mathbf{i})$, by Definition 4. Also, the failure probability averaged over the 2^m input values $\mathbf{i} \in \mathbb{Z}_2^m$ is smaller than or equal to the maximal failure probability, $\max_{\mathbf{i}\in\mathbb{Z}_2^m} p_{\text{fail}}(\mathbf{i}) \geqslant 2^{-m} \sum_{\mathbf{i}\in\mathbb{Z}_2^m} p_{\text{fail}}(\mathbf{i})$. Combining the last three relations, we find

$$2^m(1 - p_S) \geqslant p\nu.$$

To maintain contextuality, p must be bounded away from unity. With the last relation, this is guaranteed if $p_S > 1 - \frac{\nu}{2^m}$. ■

Proof of Theorem 3. For any nonlinear Boolean function, $\nu \geqslant 1$. Theorem 3 now follows from Lemma 1 with the choice $\nu = 1$. ■

If we consider all non-linear Boolean functions, no ν larger than 1 can be chosen for Theorem 3, since $o = \prod_{k=1}^m i_k$ has $\nu = 1$. Therefore, the contextuality threshold of $p_{S,\text{th}} = 1 - 1/2^m$ stated in Theorem 3 is optimal for the evaluation of general nonlinear functions. However, if we restrict to special functions, then the range of p_S for which the computation remains contextual can be significantly extended. In this regard, we recall from MacWilliams and Sloane [25] the following.

Definition 5. A Boolean function $f(v_1, \ldots, v_m)$, for m even, is called "bent" if the Hadamard transform coefficients $\hat{F}(\mathbf{u})$ given by $\hat{F}(\mathbf{u}) = \sum_{\mathbf{v}\in\mathbb{Z}_2^m} (-1)^{\mathbf{u}\cdot\mathbf{v}+f(\mathbf{v})}$ are all $\pm 2^{m/2}$.

For bent functions we note the following [25].

Theorem 4. A bent function $f(v_1, \ldots, v_m)$ is farther away from any linear function,

$$a_0 + \sum_{i=1}^m a_i v_i,$$

than any other Boolean function. More precisely, $f(v_1, \ldots, v_m)$ is bent iff the corresponding vector $\mathbf{f}$ has a distance $2^{m-1} \pm 2^{m/2-1}$ from every code word of the Reed-Muller code $\mathcal{R}(1,m)$. If f is not bent, then $\mathbf{f}$ has a distance less than $2^{m-1} - 2^{m/2-1}$ from some code word of $\mathcal{R}(1,m)$.

Using this result, we obtain the following.

Theorem 5. Let M_{p_S} be a l2-MBQC that evaluates with success probability p_S a bent function on an even number m of bits. Then M_{p_S} is contextual if $p_S > \frac{1}{2} + (\frac{1}{2})^{m/2+1}$.

Proof of Theorem 5. With Theorem 4, we may choose $\nu = 2^{m-1} - 2^{m/2-1}$, and the result follows directly from Lemma 1. ■

The low threshold of p_S in Theorem 5 is worth of note. Consider the special case of a single output bit (which for

any l2-MBQC can be obtained by discarding the other output qubits). Then, for large values of m, the output of M_{p_s} can be very close to completely random, and yet M_{p_S} remains contextual.

IV. TESTS OF CONTEXTUALITY

It is natural to ask whether Lemma 1 and Theorems 3 and 5 lead to experimental tests of contextuality. They do: one may measure a sufficiently accurate estimate of the success probability p_S of function evaluation and then compare it with the threshold in the above theorems.

But the efficiency of the measurement procedure needs to be examined. There are two potential sources of inefficiency. First, the general contextuality threshold as stated in Theorem 3 is, for large m, very close to unity. The experiment may therefore need to be repeated a very large number of times. Second, the contextuality threshold involves the *worst-case* success probability p_S, i.e., the lowest success probability over all input values. Finding the worst value of input is in general exponentially inefficient in the number m of imput bits.

Regarding the first point, we observe that the contextuality threshold depends on the evaluated function; cf. Lemma 1. While in the worst case the threshold is indeed exponentially (in m) close to unity, for many Boolean functions it is substantially lower. The bent functions are one example.

Regarding the second point, we note that the exponential inefficiency in m can be eliminated by considering the *average* success probability of function evaluation rather than the worst-case success probability.

Definition 6. A procedure τ probabilistically evaluates a vector-valued Boolean function $\mathbf{o}(\mathbf{i})$ on an m-bit input $\mathbf{i} \in \mathbb{Z}_2^m$ with average success probability $\overline{p}_S$ if

$$\frac{1}{2^m} \sum_{\mathbf{i}\in\mathbb{Z}_2^m} \mathrm{Prob}(\tau(\mathbf{i}) = \mathbf{o}(\mathbf{i})) = \overline{p}_S.$$

While, in operational terms, the average success probability $\overline{p}_S$ may be a weaker notion than the worst-case success probability p_S, it simplifies the experimental verification of contextuality. In this regard, Lemma 1 has a counterpart,

Lemma 2. Let M_{p_S} be a l2-MBQC that evaluates with average success probability $\overline{p}_S$ a vector-valued Boolean function on m input bits with distance ν to the closest linear function. If $\overline{p}_S > 1 - \frac{\nu}{2^m}$, then M_{p_S} is contextual.

Proof of Lemma 2. Equation (13) is combined with $1 - \overline{p}_S = (\sum_{\mathbf{i}\in\mathbb{Z}_2^m} p_{\mathrm{fail}}(\mathbf{i}))/2^m$, and p is bounded as before. ■

Counterparts of Theorems 3 and 5 follow from Lemma 2. The average success probability $\overline{p}_S$ of function evaluation can now be obtained by uniformly sampling from the input values, and the exponential inefficiency in m is thus removed.

Finally, we point out a relation of Lemma 2 to (non-)contextuality inequalities existing in the literature. Inverting Lemma 2, we find that if a l2-MBQC is noncontextual, then

$$\sum_{\mathbf{i}\in\mathbb{Z}_2^m} p_{\mathrm{succ}}(\mathbf{i}) \leqslant 2^m - \nu. \tag{14}$$

Here, $p_{\mathrm{succ}}(\mathbf{i})$ is the probability that the function evaluation succeeds for input $\mathbf{i}$, and 2^m is the number of measurement contexts. For the (hardest) case of $\nu = 1$, we recognize this as the so-called *logical Bell inequality* described in [26]; also see Sec. IA in [27]. Quantum-mechanical systems can violate this inequality maximally, $\sum_{\mathbf{i}\in\mathbb{Z}_2^m} p_{\mathrm{succ}}(\mathbf{i}) = 2^m$. In the setting of l2-MBQC, this maximal violation corresponds to deterministic function evaluation.

Remark 2. Inequality (14) is a constraint imposed by noncontextuality, not necessarily locality. For l2-MBQC, it can be interpreted in terms of locality in the special case of temporally flat computations.

Remark 3. Inequality (14) has been called "logical," because it arises as a sole consequence of mutually incompatible propositions. This incompatibility leads to the inviability of noncontextual HVM descriptions. In the present setting of l2-MBQC, "logical" acquires an additional meaning: the above inequality constrains the success probability of noncontextual measurement-based computations.

V. CONCLUSION

In summary, we have shown that l2-MBQCs cannot be described by no-contextual hidden-variable models if they compute nonlinear Boolean functions with a sufficiently high probability of success. The probability threshold depends on the Boolean function in question. It is very close to 1 for products of high degree but close to $1/2$ for bent functions.

In addition, we would like to draw attention to the following point. Although we have stated Theorems 2, 3, and 5 for measurement-based *quantum* computations, they hold in more general scenarios than quantum theory. An example of such a more general (and hypothetical) scenario is Popescu-Rohrlich boxes, which violate the CHSH inequality maximally. Neither the definitions [23] of contextuality applied here nor the proofs of Lemma 1 and Theorems 2, 3, and 5 use properties of quantum mechanics. Required are the binary choice for the measurement basis and a binary measurement outcome for each party, the linear processing relations Eqs. (3) and (4), and runnability (property 5 in Definition 1).

ACKNOWLEDGMENTS

The author thanks Pradeep Sarvepalli and Tzu-Chieh Wei for discussions. This work was funded by NSERC, Cifar, PIMS, and IARPA. Part of this work was performed at the Galileo Galilei Institute in Florence, Italy, during the workshop "New quantum states of matter in and out of equilibrium".

APPENDIX: LINEAR CLASSICAL PROCESSING RELATIONS IN MBQC

Since the linear relations of classical side processing in MBQC [15] are essential for our argument, we briefly review here how they come about. We discuss this for the example of the resource state being a three-qubit cluster state [see Fig. 2(a)]. The argument for general cluster and graph states [28–30] is analogous. The measured local observables are $\cos 2\phi_i\, X_i + \sin 2\phi_i\, Y_i$. This MBQC can be used to simulate the circuit shown in Fig. 2(b), consisting of the application

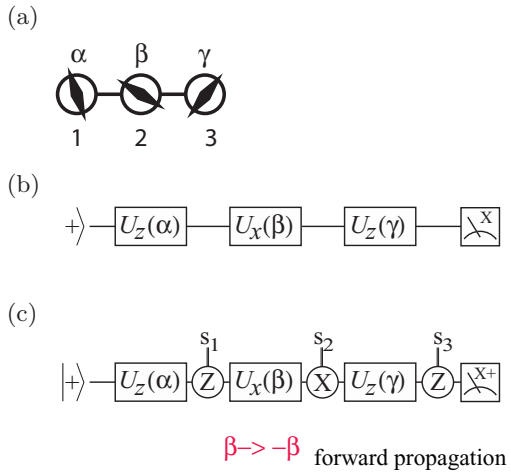

FIG. 2. (Color online) Origin of the linear processing relations (by example). (a) MBQC on a three-particle cluster state. It can be used to simulate the circuit shown in (b) but—if run straightforwardly—executes the probabilistic circuit (c). The need to compensate for the random measurement outcomes enforces a temporal order among the measurements. For explanation see text.

of a general one-qubit rotation (in its Euler decomposition) to a state $|+\rangle \sim |0\rangle + |1\rangle$, followed by a measurement of the Pauli observable X. However, if all three measurements are performed simultaneously, with measurement angles $\phi_1 = \alpha$, $\phi_2 = \beta$, and $\phi_3 = \gamma$, then, rather than the desired circuit in Fig. 2(b), the probabilistic circuit in Fig. 2(c) is realized. It differs from the desired circuit by the insertion of Pauli spin or phase flips which are conditioned on measurement outcomes obtained. These random flips need to be removed from the circuit. This task can be accomplished by measuring the three qubits in sequence and adjusting measurement bases on the go.

Consider the phase flip $(Z)^{s_1}$ next to the z rotation $U_z(\alpha)$ in Fig. 2(a). It can be propagated forward through the computation, past the readout measurement. Due to the anticommutation relation $ZX = -XZ$, on its course the Pauli operator Z flips the rotation angle β and the outcome o of the readout measurement. If qubit 1 is measured before qubit 2, the conditional flip of the rotation angle β can still be accommodated by a conditional flip of the measurement angle ϕ_2, namely, $\phi_2 = (-1)^{s_1}\beta$. The two other probabilistic insertions of Pauli flips propagate in a similar fashion. The net result is that the sign $(-1)^{q_2}$ of the measurement angle ϕ_2 is given by $q_2 = s_1$, and similarly, $q_3 = s_2$. Furthermore, the output o is $o = s_1 \oplus s_3$. These relations are all mod 2 linear. This property is a consequence of the (anti-)commutation relations of Pauli operators and generalizes to the universal case [15].

[1] P. Shor, *Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM* (IEEE Press, New York, 1994), p. 124.

[2] L. K. Grover, in *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing (STOC)* (Association for Computing Machinery, New York, 1996), p. 212.

[3] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, in *Proceedings of the 35th Symposium on the Theory of Computing* (Association for Computing Machinery, New York, 2003), p. 59.

[4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2007).

[5] R. Cleve, A. Ekert, C. Macciavello, and M. Mosca, Proc. R. Soc. London, Ser. A **454**, 339 (1998).

[6] D. Poulin, A. Qarry, R. Somma, and F. Verstraete, Phys. Rev. Lett. **106**, 170501 (2011).

[7] D. Gottesman, Ph.D. thesis, California Institute of Technology, 1997.

[8] M. van den Nest, Phys. Rev. Lett. **110**, 060504 (2013).

[9] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).

[10] M. van den Nest, W. Dür, G. Vidal, and H. J. Briegel, Phys. Rev. A **75**, 012337 (2007).

[11] M. J. Bremner, C. Mora, and A. Winter, Phys. Rev. Lett. **102**, 190502 (2009).

[12] D. Gross, S. T. Flammia, and J. Eisert, Phys. Rev. Lett. **102**, 190501 (2009).

[13] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).

[14] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[15] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[16] J. Anders and D. E. Browne, Phys. Rev. Lett. **102**, 050502 (2009).

[17] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).

[18] M. Mosca and C. Zalka, arXiv:quant-ph/0301093.

[19] M. A. Nielsen, Phys. Lett. A **308**, 96 (2003).

[20] D. W. Leung, arXiv:quant-ph/0111122; Int. J. Quantum Inform. **2**, 33 (2004).

[21] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer, Dordrecht, 1989), p. 69.

[22] D. Gross, J. Eisert, N. Schuch, and D. Perez-Garcia, Phys. Rev. A **76**, 052315 (2007).

[23] S. Abramsky and A. Brandenburger, New J. Phys. **13**, 113036 (2011).

[24] M. J. Hoban and D. E. Browne, Phys. Rev. Lett. **107**, 120402 (2011).

[25] F. J. MacWilliams and N. A. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1977).

[26] L. Hardy, Phys. Lett. A **161**, 21 (1991).

[27] S. Abramsky and L. Hardy, Phys. Rev. A **85**, 062114 (2012).

[28] H. J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).

[29] D. Schlingemann and R. F. Werner, Phys. Rev. A **65**, 012308 (2002).

[30] M. Hein, J. Eisert, and H. J. Briegel, Phys. Rev. A **69**, 062311 (2004).