Preparation Contextuality Powers Parity-Oblivious Multiplexing

Robert W. Spekkens,[1] D. H. Buzacott,[2,3] A. J. Keehn,[2,3] Ben Toner,[4] and G. J. Pryde[2,3]

[1]*DAMTP, University of Cambridge, Cambridge, CB3 0WA, United Kingdom*

[2]*Centre for Quantum Computer Technology, Griffith University, Brisbane 4111, Australia*

[3]*Centre for Quantum Dynamics, Griffith University, Brisbane 4111, Australia*

[4]*Centrum voor Wiskunde en Informatica, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands*

(Received 12 May 2008; published 5 January 2009)

In a noncontextual hidden variable model of quantum theory, hidden variables determine the outcomes of every measurement in a manner that is independent of how the measurement is implemented. Using a generalization of this notion to arbitrary operational theories and to preparation procedures, we demonstrate that a particular two-party information-processing task, "parity-oblivious multiplexing," is powered by contextuality in the sense that there is a limit to how well any theory described by a noncontextual hidden variable model can perform. This bound constitutes a "noncontextuality inequality" that is violated by quantum theory. We report an experimental violation of this inequality in good agreement with the quantum predictions. The experimental results also provide the first demonstration of 2-to-1 and 3-to-1 quantum random access codes.

DOI: 10.1103/PhysRevLett.102.010401 PACS numbers: 03.65.Ta, 03.67.−a, 42.50.Dv, 42.50.Ex

The Bell-Kochen-Specker theorem [1] shows that the predictions of quantum theory are inconsistent with a hidden variable model having the following feature: if A, B, and C are Hermitian operators such that A and B commute, A and C commute, but B and C do not commute, then the value predicted to occur in a measurement of A does not depend on whether B or C was measured simultaneously. This feature is called "noncontextuality." Significantly, it is only well-defined for models of quantum theory (and then only for projective measurements and deterministic models) [2]. By contrast, Bell's definition of a *local* model applies to any theory that can be described operationally [3]. Consequently, whereas one can test whether or not experimental statistics are consistent with a local model (by testing whether or not they satisfy Bell inequalities), there is no way to test whether or not experimental statistics are consistent with a noncontextual model (and no way of defining associated "noncontextuality inequalities") *unless* one generalizes the traditional notion of noncontextuality in such a way that it makes no reference to the quantum formalism. Suggestions for such a formulation have been made by several authors [4]. A particularly natural generalization (and slight modification) which applies to all models (deterministic or not) of any operational theory has been proposed in Ref. [2]. We here derive a noncontextuality (NC) inequality based on this notion.

Because information-theoretic tasks can be characterized entirely in terms of experimental statistics, one can explore whether theories that violate NC inequalities may provide information-theoretic advantages over theories that satisfy these inequalities. We prove that this is indeed the case for a task which we call *parity-oblivious multiplexing*, a kind of two-party secure computation. (The notion that contextuality might yield an advantage for multiplexing tasks was first put forward by Galvão [5].) The NC inequality we derive provides a bound on the probability of success in this task, and we demonstrate a quantum protocol for parity-oblivious multiplexing for which the probability of success exceeds the noncontextual bound.

Finally, we report an experimental implementation of this protocol that achieves a probability of success in good agreement with the quantum result and in violation of the NC inequality.

Operational theories and noncontextual models.—In an operational theory, the primitives of description are preparations and measurements, specified as instructions for what to do in the laboratory. The theory simply provides an algorithm for calculating the probability $p(k|P, M)$ of an outcome k of measurement M given a preparation P. As an example, in quantum theory, every preparation P is represented by a density operator ρ_P, every measurement M is represented by a positive operator valued measure $\{E_{M,k}\}$, and the probability of outcome k is given by $p(k|P, M) = \mathrm{Tr}(\rho_P E_{M,k})$.

In a hidden variable model of an operational theory, a preparation procedure is assumed to prepare a system with certain properties and a measurement procedure is assumed to reveal something about those properties. The set of all variables describing the system is denoted λ. It is presumed that for every preparation P, there is a probability distribution $p(\lambda|P)$ such that implementing P causes the system to be prepared in physical state λ with probability $p(\lambda|P)$. Similarly, it is presumed that for every measurement M, there is a distribution $p(k|\lambda, M)$ such that implementing M on a system described by λ yields outcome k with probability $p(k|\lambda, M)$. For the hidden variable model to reproduce the predictions of the operational theory, it must satisfy $p(k|P, M) = \int d\lambda\, p(k|\lambda, M) p(\lambda|P)$.

A hidden variable model is *preparation noncontextual* if the following implication holds,

$$\forall M\colon\ p(k|P,M) = p(k|P',M) \Rightarrow p(\lambda|P) = p(\lambda|P'), \quad (1)$$

that is, if two preparations yield the same statistics for all possible measurements, then they are represented equivalently in the hidden variable model. Similarly, *measurement noncontextuality* is the condition that

$$\forall P\colon\ p(k|P,M) = p(k|P,M') \Rightarrow p(k|\lambda,M) = p(k|\lambda,M'), \quad (2)$$

that is, if two measurements have the same statistics for all possible preparations, then they are represented equivalently in the model. More details can be found in Ref. [2]. An NC inequality is any inequality on experimental statistics that follows from the assumption that there exists a hidden variable model that is preparation and measurement noncontextual. It is of the form $f(p(k|P_1,M_1), p(j|P_2,M_2),\ldots) \leq C$ for some function f and constant C.

Parity-oblivious multiplexing.—Suppose that Alice and Bob wish to perform the following information-processing task, which we call *n-bit parity-oblivious multiplexing*. Alice has as input an n-bit string x chosen uniformly at random from $\{0,1\}^n$. Bob has as input an integer y chosen uniformly at random from $\{1,\ldots,n\}$ and must output the bit $b = x_y$, that is, the yth bit of Alice's input. Alice can send a system to Bob encoding information about her input; however, there is a cryptographic constraint: no information about any parity of x can be transmitted to Bob. More specifically, letting $s \in \mathrm{Par}$ where $\mathrm{Par} \equiv \{r | r \in \{0,1\}^n, \sum_i r_i \geq 2\}$ is the set of n-bit strings with at least 2 bits that are 1, no information about $x \cdot s = \bigoplus_i x_i s_i$ (termed the *s-parity*) for any such s can be transmitted to Bob (here $\oplus$ denotes sum modulo 2). This task is similar to an n-to-1 quantum random access code [5–8] except that it has a constraint of parity obliviousness rather than a constraint on the potential information-carrying capacity of the system used.

Lemma 1.—Classically, the optimal probability of success in n-bit parity-oblivious multiplexing satisfies $p(b = x_y) \leq (n+1)/2n$.

Proof.—(For details, see [9].) The only classical encodings of x that reveal no information about any parity (while encoding *some* information about x) are those that encode only a single bit x_i for some i. Given that y is uniformly distributed, it makes no difference which bit it is. Therefore, we may assume that Alice and Bob agree that Alice will always encode the first bit, x_1. If $y = 1$, which occurs with probability $1/n$, then Bob can output $b = x_y$ and win. With probability $(n-1)/n$, we have $y \neq 1$, and in this case, Bob can at best guess the value of x_y and wins with probability $1/2$. ∎

What is the most general protocol that can be implemented in an arbitrary operational theory? For each input string x, Alice implements a preparation procedure P_x, and for each integer y, Bob implements a binary-outcome measurement M_y, and reports the outcome b as his output. The probability of winning is

$$p(b = x_y) = \frac{1}{2^n n} \sum_{y\in\{1,\ldots,n\}} \sum_{x\in\{0,1\}^n} p(b = x_y | P_x, M_y) \quad (3)$$

where $1/2^n n$ is the prior probability for a particular x and y. The parity-oblivious constraint requires that for every s-parity, there is no outcome of any measurement for which posterior probabilities for s-parity 0 and s-parity 1 are different, that is,

$$\forall s\ \forall M\ \forall k\colon \sum_{x|x\cdot s=0} p(P_x|k,M) = \sum_{x|x\cdot s=1} p(P_x|k,M). \quad (4)$$

Noncontextuality inequality.—The main theoretical result of this Letter is the following theorem.

Theorem 2.—In an operational theory that admits a preparation noncontextual hidden variable model, the optimal probability of success in n-bit parity-oblivious multiplexing satisfies $p(b = x_y) \leq (n+1)/2n$.

Proof.—Define $P_{s,b}$ to be the procedure obtained by choosing uniformly at random an x such that $x \cdot s = b$ and implementing P_x. Clearly, for any measurement M, the probability of outcome k given preparation $P_{s,b}$ is simply

$$p(k|P_{s,b},M) = \frac{1}{2^{n-1}} \sum_{x|x\cdot s=b} p(k|P_x,M). \quad (5)$$

Similarly, the probability of hidden variable λ given an implementation of $P_{s,b}$ is simply

$$p(\lambda|P_{s,b}) = \frac{1}{2^{n-1}} \sum_{x|x\cdot s=b} p(\lambda|P_x). \quad (6)$$

Now note that one can reexpress the parity-oblivious condition, Eq. (4), as $\forall s\ \forall M\colon \sum_{x|x\cdot s=0} p(k|P_x,M) = \sum_{x|x\cdot s=1} p(k|P_x,M)$ (it follows from Bayes' rule and the uniformity of the prior over x). Combining this with Eq. (5), we infer that $\forall s\ \forall M\colon p(k|P_{s,0},M) = p(k|P_{s,1},M)$ which is simply the statement that mixed preparations corresponding to opposite s-parities are indistinguishable by any measurement. But together with the assumption that the hidden variable model is preparation noncontextual, Eq. (1), this implies that $\forall s\colon\ p(\lambda|P_{s,0}) = p(\lambda|P_{s,1})$, which states that mixed preparations corresponding to opposite s-parities are also indistinguishable at the hidden variable level. Using Eq. (6) and Bayes' rule again, we obtain

$$\forall s\colon \sum_{x|x\cdot s=0} p(P_x|\lambda) = \sum_{x|x\cdot s=1} p(P_x|\lambda). \quad (7)$$

Therefore, even if one knew λ, the posterior probabilities for s-parity 0 and s-parity 1 would be the same, that is, one would know nothing about any s-parity of x. The argument so far can be summarized as follows: for preparation noncontextual models, parity obliviousness at the operational level implies parity obliviousness at the level of the hidden variables.

The hidden state λ provides a classical encoding of x. But, as just shown, it is one that cannot contain information about any s-parity. We recall from lemma 1 that such encodings have at most 1 bit of information, x_i, about x. Consequently, even if Bob could determine λ perfectly, he and Alice could at best achieve the optimal probability of success achievable in a classical protocol (specified in lemma 1), while if Bob is limited in his ability to determine λ (as will be the case in general in a hidden variable model), they will do worse. ■

*Quantum case.—*We now consider how well one can achieve parity-oblivious multiplexing in quantum theory. The following is a protocol for the 2-bit case that uses a single qubit as the quantum message. Alice encodes her 2 bits into the four pure quantum states with Bloch vectors $(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$ equally distributed on an equatorial plane of the Bloch sphere, as indicated in Fig. 1 [recall that a density operator ρ is related to its Bloch vector $\vec{r}$ by $\rho = \frac{1}{2}(I + \vec{r}\cdot\vec{\sigma})$, where $\vec{\sigma}$ is the vector of Pauli matrices]. Bob measures along the $\hat{x}$ axis if he wishes to learn the first bit, and along the $\hat{y}$ axis if he wishes to learn the second. He guesses the bit value 0 upon obtaining the positive outcome. In all cases, the guessed value is correct with probability $\cos^2(\pi/8) \simeq 0.853553$. Meanwhile, no information about the parity can be obtained by any quantum measurement given that the parity 0 and parity 1 mixtures are represented by the same density operator, $\frac{1}{2}\rho_{00} + \frac{1}{2}\rho_{11} = \frac{1}{2}\rho_{01} + \frac{1}{2}\rho_{10} = I/2$. We have a violation of the NC inequality of Thm. 2 because for $n = 2$, the upper bound on the probability of success is $3/4$. By exploiting a connection with the Clauser-Horne-Shimony-Holt inequality [10], one can show that this protocol yields the maximum possible quantum violation of the NC inequality.

A protocol for 3-bit parity-oblivious multiplexing using a single qubit proceeds as follows. Alice encodes her 3 bits into a set of eight pure quantum states associated with Bloch vectors $(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})$ forming a cube inside the Bloch sphere (see Fig. 1). Bob measures along the $\hat{x}$, $\hat{y}$, or $\hat{z}$ axes to obtain the first, second, or third bits. In all cases, the guessed value is correct with probability $\frac{1}{2}\times(1+\frac{1}{\sqrt{3}}) \simeq 0.788675$. The mixture of the four states corresponding to $x_1 \oplus x_2 = 0$ [i.e., s-parity 0 for $s = (1, 1, 0)$] is identical to the mixture of the four states corresponding to $x_1 \oplus x_2 = 1$ and is equal to $I/2$. Similarly for the two mixtures associated with each of the other three parities, $x_1 \oplus x_3$ [$s = (1, 0, 1)$], $x_2 \oplus x_3$ [$s = (0, 1, 1)$], and $x_1 \oplus x_2 \oplus x_3$ [$s = (1, 1, 1)$]. The protocol is therefore parity oblivious for all s-parities. Again, we have a violation of the NC inequality because for $n = 3$, the upper bound on the probability of success is $2/3$. It is an open question whether 0.788675 is the maximum possible quantum violation.

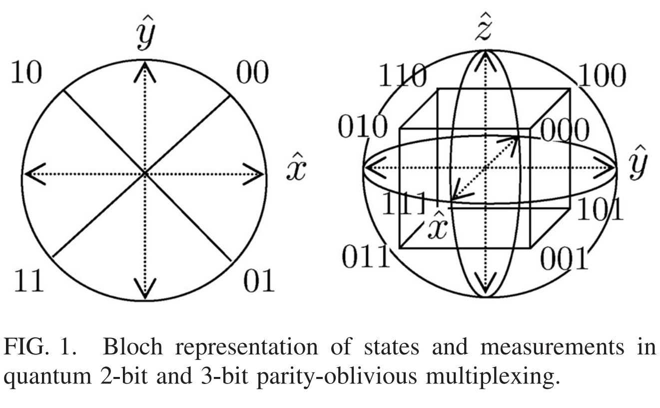

FIG. 1. Bloch representation of states and measurements in quantum 2-bit and 3-bit parity-oblivious multiplexing.

The 2-bit protocol was originally presented as a 2-to-1 quantum random access code by Wiesner [6] and rediscovered in Ref. [7], while the 3-bit protocol was presented in Ref. [8] as an instance of a 3-to-1 quantum random access code (the original idea is attributed to Chuang in Ref. [7]).

*Experimental results.—*We experimentally demonstrate better-than-classical performance for 2-bit and 3-bit parity-oblivious multiplexing by implementing the quantum protocols using polarization qubits. Photon pairs from down-conversion are coupled into single mode optical fibers. One photon acts as a trigger, while the other is used in the experiment. Alice's state preparation consists of a fiber polarization controller, and a polarizing beam displacer, rotated to the input state angle, used to ensure high-purity linearly polarized states for the 2-bit protocol. An additional quarter wave plate is used to prepare elliptically polarized states for the 3-bit protocol. Bob's measurement consists of a polarizing beam displacer mounted in a computer-controlled rotation mount, followed by a single photon counting module. For our demonstration, a detector is placed at only a single output port of the beam displacer and the probability of each outcome is calculated from the relative number of counts for a given beam displacer angle and the one orthogonal to it [9]. Adjustment of the beam displacer and quarter wave plate angles allows measurement of the horizontal/vertical basis, the diagonal-antidiagonal basis, and the right/left-circular basis. Valid measurement events are heralded by a coincidence count between the directly detected photon and the experiment photon. These experimental procedures for a given x and y define the preparation P_x and the measurement M_y, respectively.

We obtained probabilities $p(k = x_y|P_x, M_y)$ by accumulating statistics over approximately 3.5×10^7 coincidence counts for each x and y in the 2-bit scheme and 2.4×10^7 in the 3-bit scheme. Using Eq. (3), we calculated the 2-bit and 3-bit probabilities of success to be $p(b = x_y) = 0.851929 \pm 0.000030$ and $p(b = x_y) = 0.786476 \pm 0.000017$, respectively. The errors were determined from the Poissonian counting statistics of the parametric source and the small repeatability error in the wave plate settings, using standard error analysis techniques. These probabilities of success violate the NC inequality of Thm. 2 with a high degree of confidence: 3410 and 6922 standard deviations, respectively. They are also close to the predicted quantum values of 0.853553 and 0.788675, achieving a violation that is 98.4% and 98.2%, respectively, of the gap between the NC bound and the quantum value.

Just as Bell inequality violations are only surprising given the absence of signalling between the two wings of the experiment, the NC inequality violations are only surprising given the parity-oblivious property. However, whereas one can establish the absence of signalling by confirming that the two wings are spacelike separated, one must directly test for transmission of information about the parity in our experiment. A consideration of how this is to be accomplished highlights two shortcomings in the operational definition of preparation noncontextuality of Eq. (1): in practice, one can never implement *all* measurements and one never finds truly *identical* statistics. The first issue may be addressed by relying on previous experimental evidence for the existence of a tomographically complete set of measurements—one from which the statistics of any other measurement can be calculated—and testing indistinguishability relative to this set alone, as we shall do here. The second issue may be addressed by presuming a kind of continuity: closeness of experimental statistics implies closeness of the representations in the model [2] (this parallels the problem of dealing with imperfect alignment in traditional proofs of contextuality [11], where continuity also provides an answer [4,12]). In the present work, we simply demonstrate that the experimental statistics are *close* to parity oblivious while yielding a large violation of the noncontextuality inequalities, and leave a more detailed analysis for future work.

We quantify the obliviousness of our experimental protocol for a particular s-parity by the maximum probability that Bob can correctly estimate this parity in a variation over all measurements. One can estimate this by implementing a tomographically complete set of measurements, then reconstructing the states ρ_0 and ρ_1 associated with s-parity 0 and s-parity 1, and finally making use of the fact that the maximum probability of discriminating these states is $\frac{1}{2} + \frac{1}{4}\,\mathrm{Tr}|\rho_0 - \rho_1|$. Among all s-parities, we calculate the largest such probability to be 0.5020 ± 0.0002. This calculation is not sufficient, however, because it neglects an imperfection in the experiment that also contributes to leakage of information about the parity, namely, that there is a small probability of more than one photon being sent to the experiment. By our characterization of the source, we estimate the probability of two photons to be 0.007 ± 0.003 relative to the single photon generation probability. If two photons pass through the polarizers in the ideal protocol, the maximum probability of correctly estimating the parity can be quite far from $1/2$: it is $3/4$ in the case of the 2-bit scheme and $2/3$ for three of the four s-parities in the 3-bit scheme. However, the fact that this possibility occurs with low probability implies that the two-photon contribution to the probability of correct estimation is comparable to the one-photon contribution. (Contributions from three or more photons are negligible in comparison). The weighted average of these contributions is easily calculated and the largest, among all s-parities, is found to be 0.504 ± 0.002. The fact that this is within 1% of $1/2$ demonstrates that our experimental protocols are indeed close to parity oblivious.

Given that the quantum protocols described herein are also 2-to-1 and 3-to-1 random access codes, our results constitute the first experimental demonstration of a quantum advantage for these tasks as well.

Finally, it is worth noting that every Bell inequality is a special case of an NC inequality where all assumptions of noncontextuality are justified by locality [2]. Consequently, every experimental violation of a Bell inequality demonstrates the impossibility of a noncontextual hidden variable model. Indeed, this is *all* that can be demonstrated by those that fail to seal the locality loophole [13,14]. Nonetheless, a dedicated experiment of the sort we have described here can achieve a large violation with high confidence at a smaller cost of experimental effort.

R. W. S. thanks M. Leifer and J. Barrett for helpful discussions. This work has been supported by the Australian Research Council, an IARPA-funded US Army Research Office contract, NWO VICI Project 639-023-302, the Dutch BSIK/BRICKS project, the EUs FP6-FET Integrated Projects SCALA and QAP, and the Royal Society.

[1] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966); S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[2] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[3] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[4] A. Cabello and G. Garcia-Alcaine, Phys. Rev. Lett. **80**, 1797 (1998); C. Simon, C. Brukner, and A. Zeilinger, Phys. Rev. Lett. **86**, 4427 (2001); J.-A. Larsson, Europhys. Lett. **58**, 799 (2002).
[5] E. F. Galvão, Ph.D. thesis, arXiv:quant-ph/0212124v1.
[6] S. Wiesner, SIGACT News **15**, 78 (1983).
[7] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, in *Proceedings of the 31st Annual ACM Symposium on the Theory of Computing* (ACM Press, New York, 1999).
[8] M. Hayashi *et al.*, New J. Phys. **8**, 129 (2006).
[9] See EPAPS Document No. E-PRLTAO-102-024902 for a detailed proof of the optimal probability of success for classical n-bit parity-oblivious multiplexing and further details of the experimental setup. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[10] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[11] D. A. Meyer, Phys. Rev. Lett. **83**, 3751 (1999).
[12] N. D. Mermin, arXiv:quant-ph/9912081.
[13] M. A. Rowe *et al.*, Nature (London) **409**, 791 (2001).
[14] Y. Hasegawa *et al.*, Nature (London) **425**, 45 (2003).