Tight Limits on Nonlocality from Nontrivial Communication Complexity; a.k.a. Reliable Computation with Asymmetric Gate Noise

Noah Shutty*†‡, Mary Wootters*§, and Patrick Hayden*†¶
*{noaj,marykw,phayden}@stanford.edu
†Stanford Institute for Theoretical Physics, Stanford University, Stanford California 94305, USA
§Departments of Computer Science and Electrical Engineering, Stanford University, Stanford California 94305, USA. M.W. was supported in part by NSF CAREER CCF-1844628.
‡N.S. was supported in part by NSF DGE-1656518.
¶P.H. was supported by AFOSR (FA9550-16-1-0082), CIFAR and the Simons Foundation.

***Abstract*—It has long been known that the existence of certain superquantum nonlocal correlations would cause communication complexity to collapse. The absurdity of a world in which any function could be evaluated by two players with a constant amount of communication in turn provides a tantalizing way to distinguish quantum mechanics from incorrect theories of physics; the statement "communication complexity is nontrivial" has even been conjectured to be a concise information-theoretic axiom for characterizing quantum mechanics. We directly address the viability of that perspective with two results. First, we exhibit a nonlocal game such that communication complexity collapses in any physical theory whose maximal winning probability exceeds the quantum value. Second, we consider the venerable CHSH game that initiated this line of inquiry. In that case, the quantum value is about 0.85 but it is known that a winning probability of approximately 0.91 would collapse communication complexity. We show that the 0.91 result is the best possible using a large class of proof strategies, suggesting that the communication complexity axiom is insufficient for characterizing CHSH correlations. Both results build on new insights about reliable classical computation. The first exploits our formalization of an equivalence between amplification and reliable computation, while the second follows from a rigorous determination of the threshold for reliable computation with formulas of noise-free XOR gates and noisy AND gates.**

***Keywords*-Quantum entanglement; Communication Complexity, Fault-Tolerant Computing; Computational complexity; Logic gates; Probabilistic computing; Protocols;**

I. Introduction

A proposed information-theoretic axiom for characterizing quantum mechanics is that "communication complexity is nontrivial". That is, two parties with inputs $x \in \{0,1\}^n$ and $y \in \{0,1\}^n$ respectively should not be able to compute arbitrary functions $f(x,y)$ with high probability, using only a constant amount of communication (independent of n). Such a requirement is satisfied by quantum mechanics, and is also known to rule out superquantum success at certain *nonlocal games*. For example, consider the famous *CHSH game*.[1] The two players, Alice and Bob, cannot communicate. Alice and Bob receive independent random bits x and y respectively. Their goal is to output bits a and b, respectively, so that $a \oplus b = x \wedge y$.

In a classical world, Alice and Bob can win the CHSH game with probability $3/4$ (*e.g.* by outputting $a, b = 0$) and cannot do any better; thus the classical value of the CHSH game is $\omega_C(CHSH) = \frac{3}{4}$. If Alice and Bob have access to any *nonsignalling* correlation—that is, they can produce correlated bits a and b in any way they like as long as they do not gain the ability to communicate—then they can win the CHSH game with probability 1; we say that the nonsignalling value of the CHSH game is $\omega_{NS}(CHSH) = 1$. If instead Alice and Bob share quantum entanglement, they can do something in between ω_C and

[1] This game is named after Clauser, Horne, Shimony, and Holt and was introduced implicitly in their paper [1].

2575-8454/20/$31.00 ©2020 IEEE
DOI 10.1109/FOCS46700.2020.00028

ω_{NS}: it turns out that the quantum value of the CHSH game is [2]

$$\omega_Q(CHSH) = \frac{1}{2} + \frac{1}{\sqrt{8}} \approx 0.8536.$$

Work of van Dam [3], and independently Cleve [4], showed that if Alice and Bob could win the CHSH game with probability 1, then communication complexity would become trivial. This was extended by Brassard *et al.* [5], who showed that if Alice and Bob could win the CHSH game with probability greater than $\frac{1}{2} + \frac{1}{\sqrt{6}} \approx 0.908$ then communication complexity would become trivial. Thus, the axiom "communication complexity is nontrivial" in some sense explains why $\omega_Q(CHSH) < 0.908$. Other works have extended the set of nonlocal correlations known to collapse communication complexity [6], [7], [8].

However, so far the axiom "communication complexity is nontrivial" had not pinned down the exact quantum value for any nonlocal game. For example, in the CHSH game, there is a gap between the threshold of approximately 0.908 that Brassard *et al.* obtain and the true quantum value $\omega_Q(CHSH) \approx 0.853$.

In this paper, we address this question: can the axiom "communication complexity is nontrivial" be used to explain the quantum value of certain nonlocal games? Along the way, we formalize a connection to the theory of reliable computation for (classical) circuits with noisy gates, and our results for nonlocal games correspond to new results for reliable classical computation. We outline our contributions in both areas below.

A. Contributions

First, we address the extent to which the axiom "communication complexity is nontrivial" can explain the quantum value of nonlocal games.

(1) We exhibit a nonlocal game G, for which

$$\omega_C(G) < \omega_Q(G) < \omega_{NS}(G),$$

and for which the axiom "communication complexity is not trivial" precisely pins down the value $\omega_Q(G)$. Our game G is "complete", in the sense that if communication complexity is trivial in any superquantum theory S, then there is (a version of) our game G so that $\omega_S(G) > \omega_Q(G)$. That is, a superquantum advantage at the game G makes communication complexity trivial and, meanwhile, any universe in which communication complexity is trivial offers a superquantum advantage at the game G.

(2) We provide evidence that the axiom "communication complexity is nontrivial" is in fact *not* sufficient to pin down the quantum value of the CHSH game itself. In more detail, in [5], Brassard *et al.* use the ability to succeed at the CHSH game essentially as a noisy AND gate. They show that reliable computation is possible when these noisy AND gates are used along with noiseless XOR gates (which correspond to certain local operations for Alice and Bob). This leads to protocols that collapse communication complexity. We show that this strategy cannot be pursued further: the threshold of 0.908 is tight for this model of computation. While this result is only a barrier against one line of attack, it does suggest that the axiom "communication complexity is nontrivial" may not suffice to explain $\omega_Q(CHSH)$.

As alluded to in our contribution (2) above, there is a connection to reliable computation with noisy gates. In that area, we make the following contributions.

(3) Our contribution (2) above can be seen as a result about reliable computation. Consider the following circuit model with noisy gates. Let $\wedge_\varepsilon$ denote a 2-input AND gate which, for any input produces an incorrect answer with probability ε, and let $\oplus_0$ denote a (noiseless) 2-input XOR gate.[2] Let $\mathcal{C}_\varepsilon$ be the collection of formulas[3] defined on the gate set $\{\wedge_\varepsilon, \oplus_0\}$, where the noise in each $\wedge_\varepsilon$ gate is independent.

[2]Here and in the rest of the paper, for a gate g, g_ε refers to a version of g which fails with probability ε.

[3]A *formula* is a circuit where every gate has fan-out 1 (that is, the graph underlying the circuit is a tree and each input variable may appear at one or more leaves of this tree).

Our main technical result is that the noise threshold for reliable computation for this model is precisely $\varepsilon = 1/6$. That is, when $\varepsilon < 1/6$, it is possible to compute any function using a formula in $\mathcal{C}_\varepsilon$ with error probability bounded away from $1/2$ for each possible input. On the other hand, for any $\varepsilon \geq 1/6$, there is some function for which this is impossible.

There has been a great deal of work on pinning down noise thresholds for reliable computation, which we survey in Section III. However, most prior work has focused on *symmetric noise,* where the noise rate is the same across all gate types. As we discuss below in Section II, extending these results to asymmetric noise—and in particular to include noiseless gates—raises several new challenges. Table I summarizes how our work compares with prior results, and this is discussed further in Section III.

Beyond our primary motivation in quantum mechanics, we believe that the case of asymmetric gate noise is an independently interesting direction in fault-tolerant computation. We hope that our techniques and results may spur future research in this direction.

(4) We formalize an equivalence between reliable computation and *amplification.* Informally, an amplifier is a function $f : \{0,1\}^d \to \{0,1\}$ so that when f is fed in random bits $x \in \{0,1\}^d$ with a slight bias away from $1/2$, the output $f(x)$ amplifies that bias. While a relationship between reliable computation and amplification had been present in prior work, nailing down an equivalence is a bit subtle, and requires considering the *convex hull* of circuit classes; to the best of our knowledge ours is the first work to do this.

Our equivalence between reliable computation and amplification is required to establish the threshold in our contribution (3) above. Further, it leads to the definition and analysis of our game G from contribution (1) whose quantum value is pinned down by the nontriviality of communication complexity.

B. Organization

In Section II, we state our results in more detail, and give an overview of our proof techniques. In Section III, we survey related work. We defer to the complete version of this paper for all formal definitions and proofs [9]. We conclude in Section IV with some discussion and future directions.

II. Results and Technical Overview

A. Relationship between nontrivial communication complexity and reliable computation

In our study of nonlocal games, we will consider players Alice and Bob who have joint access to different sets of *bipartite correlations.* A bipartite correlation can be thought of as a box that a spatially separated Alice and Bob can use to process distributed inputs, without providing them the ability to communicate. Alice inputs x, Bob inputs y, and the box outputs a for Alice and b for Bob according to some distribution $\mathbb{P}[a, b|x, y]$. In contrast to a nonlocal game, a bipartite correlation has no assumed distribution on its inputs. That is, a bipartite correlation is simply a stochastic map. Sets of bipartite correlations of interest include C, the set of all bipartite correlations that are possible classically; Q, the set of all bipartite correlations that are possible if Alice and Bob share quantum entanglement; and NS, the set of all nonsignalling bipartite correlations.

In this paper we will consider sets S of bipartite correlations that are closed under all of the operations that Alice and Bob might want to do to combine elements of S with each other (for example, composing or taking probabilistic mixtures of correlations); following [10], we say that such sets are *closed under wirings.* The sets C, Q, and NS are all closed under wirings. A set S of bipartite correlations naturally gives rise to a circuit class by thinking about how these correlations act on *distributed bits.* We say that a bit z is distributed as $z = x \oplus y$ if Alice holds x and Bob holds y, where x is uniformly random and $y = z \oplus x$. Then we can think of a bipartite correlation acting on inputs x and y as a gate acting on input z.

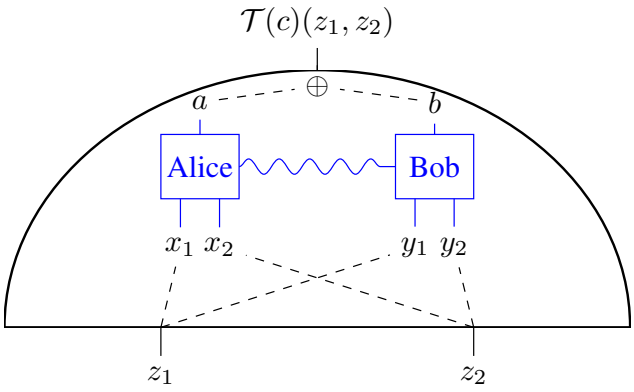

Figure 1: Defining a randomized circuit $\mathcal{T}(c)$ from a bipartite correlation $c \in S$. Here $\mathcal{T}(c)$ takes input bits $\mathbf{z} \in \{0,1\}^2$, i.e. $t = 2$ in the discussion.

In more detail, suppose that $c \in S$ is a bipartite correlation that stochastically maps inputs $\mathbf{x}, \mathbf{y} \in \{0,1\}^t$ for Alice and Bob respectively to bits $a, b \in \{0,1\}$. We can define a randomized gate $\mathcal{T}(c) : \{0,1\}^t \to \{0,1\}$ as follows. The gate $\mathcal{T}(c)$ takes as input $\mathbf{z} \in \{0,1\}^t$. Each coordinate z_i of $\mathbf{z}$ is distributed between Alice and Bob as $z_i = x_i \oplus y_i$. Then $\mathcal{T}(c)$ outputs $a \oplus b$, where a, b are the output of c acting on $\mathbf{x}$ and $\mathbf{y}$. This process is depicted in Figure 1.

Given a convex set S of bipartite correlations that is closed under wirings, we can define[4] the set $\mathcal{T}(S)$ to be the set of circuits one can make out of the gates $\{\mathcal{T}(c) : c \in S\}$.

Our goal is to understand which sets S of bipartite correlations cause communication complexity to be trivial. We will do so by studying when the (noisy) circuit model $\mathcal{T}(S)$ supports reliable computation.

Definition II.1. *A (noisy) circuit model $\mathcal{C}$* **supports reliable computation with advantage** *$\delta_0 > 0$ if for all $n > 0$, for all Boolean functions $f : \mathbb{F}_2^n \to \mathbb{F}_2$, there exists a circuit $c \in \mathcal{C}$ such that for each possible input $\mathbf{x} \in \mathbb{F}_2^n$,*

$$(-1)^{f(\mathbf{x})} \left(\mathbb{P}[c(\mathbf{x}) = 0] - \mathbb{P}[c(\mathbf{x}) = 1]\right) \geq \delta_0, \quad (1)$$

where the probability is over the randomness in c.

[4]Technically, our definition of $\mathcal{T}(S)$ takes the convex hull of circuits comprised of $\{\mathcal{T}(c) : c \in S\}$ gates; if S is closed under wirings then this distinction does not matter, as we prove in [9].

We say that $\mathcal{C}$ **supports reliable computation** *if there exists a $\delta_0 > 0$ so that $\mathcal{C}$ supports reliable computation with advantage δ_0.*

We will say that a set S of bipartite correlations causes communication complexity to become trivial if there is some way for Alice and Bob to use the correlations in S, along with shared randomness and arbitrary local computation, to compute any function with high probability with constant communication complexity. That is, there are some constants $\varepsilon > 0$ and $t \geq 1$ so that for any n and for any $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, the following holds. There is some protocol Π_S for Alice and Bob so that Π_S uses t bits of communication (in either direction), and so that for any inputs $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$ for Alice and Bob respectively,

$$\mathbb{P}[\Pi_S(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \mathbf{y})] \geq 1/2 + \varepsilon.$$

Our starting point is the observation, implicit in [5], that if $\mathcal{T}(S)$ supports reliable computation, then S causes communication complexity to become trivial. In fact, it is not hard to see that the converse is true as well. Thus, we have the following proposition.

Proposition II.2. *Suppose that $C \subseteq S \subseteq NS$ and that S is closed under wirings. Then S causes probabilistic communication complexity to become trivial (in the sense described above) if and only if $\mathcal{T}(S)$ supports reliable computation.*

The proof of Proposition II.2 follows similar logic to [5]. The basic idea behind the connection is as follows. If $\mathcal{T}(S)$ supports reliable computation, then in particular $\mathcal{T}(S)$ contains a circuit $\mathrm{Amp} : \{0,1\}^t \to \{0,1\}$ that acts as an *amplifier*. That is, given independent random bits $z_1, \ldots, z_t$ with bias $p > 1/2$ (resp. $p < 1/2$), $\mathrm{Amp}(z_1, \ldots, z_t)$ outputs a bit a that is very likely to be 1 (resp. 0).[5] Suppose that Alice and Bob want to compute some function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. It turns out that Alice and Bob can, using only classical techniques and without communicating,

[5]Indeed, Amp is simply the circuit that implements the Majority function.

obtain bits $x_1, \ldots, x_t$ and $y_1, \ldots, y_t$ so that for each i, $x_i \oplus y_i$ has a very slight bias towards the correct answer. However, this bias shrinks as n grows. To amplify their success so that this bias is a constant, Alice and Bob use S, as shown in Figure 1, to obtain bits a and b respectively so that $a \oplus b = \mathrm{Amp}(x_1 \oplus y_1, \ldots, x_t \oplus y_t)$. Then $a \oplus b$ is very likely to be equal to the correct value of f. Finally, Alice sends the single bit a to Bob, who outputs $a \oplus b$.

Due to Proposition II.2, we shall take "$\mathcal{T}(S)$ supports reliable computation" as the *definition* of trivial probabilistic communication complexity.

With the connection between trivial communication complexity and reliable computation established, we continue with an overview of our main results in both nonlocality and in reliable communication.

B. A nonlocal game whose quantum value is the threshold for nontrivial communication complexity

Our main result in this section is the following.

Theorem II.3 (A game whose quantum value is the threshold for nontrivial communication complexity)**.** *There exists a sequence of 2-player nonlocal games G_k for $k \geq 1$ that satisfies properties (1-3) below, in which S is any set of bipartite nonsignalling correlations closed under wirings and such that $S \supseteq Q$.*

1) *For all $k \geq 1$, $\omega_C(G_k) < \omega_Q(G_k) < 1$.*
2) *Fix any $k \geq 1$. If $\omega_S(G_k) > \omega_Q(G_k)$, then S has trivial probabilistic communication complexity.*
3) *If S has trivial probabilistic communication complexity, then there exists some $k \geq 1$ such that $\omega_S(G_k) > \omega_Q(G_k)$.*

The proof of Theorem II.3 is in [9], and we sketch the intuition below. In Section II-E below, we state Theorem II.8, which roughly says that reliable computation is equivalent to containing an amplifier. Inspired by this, we define the *amplification game* Amp_k as follows. Alice and Bob get $\mathbf{x}, \mathbf{y} \in \{0,1\}^{2k+1}$ respectively, and their goal is to output a, b so that

$$a \oplus b = \mathrm{Maj}(\mathbf{x} \oplus \mathbf{y}),$$

where $\oplus$ is applied coordinate-wise. The inputs $\mathbf{x}$ and $\mathbf{y}$ are drawn from a distribution so that better-than-quantum success at the amplification game using correlations in S implies that there exists an amplifier in $\mathcal{T}(S)$. Since, by Theorem II.8, amplification is equivalent to reliable computation, this translates to reliable computation for $\mathcal{T}(S)$, which in turn, via Proposition II.2, translates into trivial probabilistic communication complexity. Formalizing these connections imply that the family Amp_k satisfies properties 2 and 3 of Theorem II.3.

However, it turns out that property 1 of Theorem II.3 is not satisfied: $\omega_C(\mathrm{Amp}_k) = \omega_Q(\mathrm{Amp}_k)$. This is disappointing if the goal is to use the axiom "communication complexity is nontrivial" to pin down $\omega_Q(\mathrm{Amp}_k)$, because it also pins down $\omega_C(\mathrm{Amp}_k)$. To obtain our game G_k as in Theorem II.3, we use Amp_k along with the *Mermin-Peres magic square game* M [11], [12] in order to make a game which retains properties 2 and 3, but which also has a gap between $\omega_C(G_k)$ and $\omega_Q(G_k)$.

Remark II.4. *The only property of the game M that we use is pseudo-telepathy – that $\omega_C(M) < \omega_Q(M) = 1$. Therefore our construction is not specific to quantum mechanics and works out for any superclassical correlation set featuring pseudo-telepathy and nontrivial communication complexity. (An example of such a set, $\tilde{Q}$, is discussed in Section III-A.) Our construction demonstrates that an idea pursued by [5] works out: for certain nonlocal games with quantum advantage, the exact quantum value is determined by the axiom "communication complexity is nontrivial". However, the general applicability of our construction to all theories with the aforementioned properties perhaps fails to illuminate the nature of quantum mechanics in particular as much as one might have hoped.*

C. The approach of Brassard et al. cannot be improved

Our next result is that the approach of Brassard *et al.* in [5] cannot be improved. Recall from the introduction that [5] shows that, if $\omega_S(CHSH) > 0.908$ for some set S of bipartite correlations, then communication complexity is trivial in any world

where S is allowed. The hope would be to extend this result to replace 0.908 with $\omega_Q(CHSH) \approx 0.854$. If this were the case, then the axiom "communication complexity is nontrivial" could pin down the quantum value of the CHSH game.

Unfortunately, we show that the approach of [5] cannot be improved. As per Proposition II.2, [5] show that $\mathcal{T}(S)$ supports reliable computation for any S that allows Alice and Bob to win the CHSH game with probability greater than 0.908. Their approach is to show that for any such S, $\mathcal{T}(S)$ contains the gates $\{\wedge_\varepsilon, \oplus_0\}$ for $\varepsilon < 1/6$. Then they show how to build an amplifier as a formula on these gates.

The hope to improve the result of [5]—to replace the threshold 0.908 with a smaller number—was to make an amplifier out of $\{\wedge_\varepsilon, \oplus_0\}$ for $\varepsilon \geq 1/6$. However, our main technical result, Theorem II.5, shows that this is impossible. That is, the class of formulas on $\{\wedge_\varepsilon, \oplus_0\}$ does not support reliable computation for any $\varepsilon \geq 1/6$, and in particular (using Theorem II.8 about the equivalence between amplification and reliable computation), it does not contain an amplifier.

This rules out the approach of [5], but there are still two avenues open. First, one could consider circuits instead of formulas. Second, one could hope to use more of the class $\mathcal{T}(S)$ than just $\{\wedge_\varepsilon, \oplus_0\}$ gates. However, there are reasons to be pessimistic about both of these avenues. First, to the best of our knowledge there are no classes of gates known for which the threshold on reliable computation for circuits is different than that for formulas. Second, although we cannot currently rule it out, we would find it surprising if there were a more efficient way of using the ability to succeed at the CHSH game than to create noisy AND gates. Thus, our results suggest that the axiom "communication complexity is nontrivial" may not pin down $\omega_Q(CHSH)$.

D. Sharp thresholds for reliable computation in $\mathcal{C}_\varepsilon$

We now explain our results on reliable computation. We begin with our main technical result, which is that the noise threshold for reliable computation using formulas on $\{\wedge_\varepsilon, \oplus_0\}$ is $\varepsilon = 1/6$.

In fact, we show something stronger, in that we allow *probabilistic mixtures* of formulas. For a class of probabilistic circuits $\mathcal{C}$, we define $\operatorname{conv}\mathcal{C}$ to be the set of probabilistic circuits obtained as distributions on elements of $\mathcal{C}$. With this notation, our main theorem in this section is as follows.

Theorem II.5 (Sharp threshold for reliable computation)**.** *Let $\varepsilon \in [1/6, 5/6]$. Let $\mathcal{C}_\varepsilon$ be the class of formulas on $\{\wedge_\varepsilon, \oplus_0\}$. Then $\operatorname{conv}\mathcal{C}_\varepsilon$ does not support reliable computation.*

The work of [5] implies that $\mathcal{C}_\varepsilon$ supports reliable computation for all $\varepsilon < 1/6$. Thus, Theorem II.5 is tight. The proof of Theorem II.5 is given in [9].

Remark II.6 (NOT gates)**.** *A noise-free $\oplus_0$ gate may be used to construct a noise-free unary $\neg$ (NOT) gate, by setting one of the input wires to 1. Thus, $\mathcal{C}_\varepsilon$ also includes $\neg_0$.*

In fact, our entire proof (including Lemma II.7 below which does not include noiseless $\oplus_0$ gates) goes through in the presence of $\neg_0$ gates, and implies the slightly stronger statement that, defining the circuit model $\mathcal{C}_{\varepsilon,\tau,0}$ of formulas from the gate set $\{\wedge_\varepsilon, \oplus_\tau, \neg_0\}$, $\operatorname{conv}\mathcal{C}_{\varepsilon,\tau,0}$ does not support reliable computation for any $\varepsilon \in [1/6, 5/6]$, $\tau \in [0, 1]$.

We need to consider convex hulls for the connection to nonlocality, described in Section II-C. A noisy circuit corresponds to a strategy for the CHSH game, for which Alice and Bob are allowed shared randomness and hence can execute probabilistic mixtures of strategies.

The main ingredient in the proof of Theorem II.5 is the following lemma.

Lemma II.7. *Let $\mathcal{C}_{\varepsilon,\tau}$ be the class of formulas on $\{\wedge_\varepsilon, \oplus_\tau\}$, and suppose that $\varepsilon \in (1/6, 5/6)$, and $\tau \in (0, 1)$. Fix $\Delta > 0$ and let $f : \{0,1\}^n \to \{0,1\}$ be a function that is computable with probability at least $1/2 + \Delta$ by stochastic maps in $\operatorname{conv}\mathcal{C}_{\varepsilon,\tau}$. Then f depends on at most a constant number of inputs.*

The proof of Lemma II.7 is given in [9]. Our proof may be viewed a probabilistic analogue of an argument first presented by Pippenger, which reduces the problem of formulas reliably computing

functions that depend on many arguments to the problem of deep formulas computing a function of a single argument [13]. The proof idea is as follows. Let f be some function. We show that for any distribution on formulas $C \in \operatorname{conv} \mathcal{C}_{\varepsilon,\tau}$, there is some variable X_i that f depends on, which appears reasonably deep, on average, in the formulas in the support of C. This means that X_i must pass through many noisy gates before reaching the output, which implies that C cannot compute f too accurately. While the basic idea is similar to the argument of [13], since we consider distributions on formulas and also allow for arbitrarily small $\tau > 0$, new ideas are required to establish Lemma II.7.

Lemma II.7 comes close to establishing Theorem II.5. Indeed, since there are functions which depend on more than a constant number of inputs (for example, the AND of n bits), Lemma II.7 implies that such functions cannot be computed in $\operatorname{conv} \mathcal{C}_{\varepsilon,\tau}$ with any constant probability larger than $1/2$, provided that $\varepsilon \in (1/6, 5/6)$ and $\tau \in (0, 1)$. The final step to the proof of Theorem II.5 is to handle the case of $(\varepsilon, \tau) \in \{1/6, 5/6\} \times \{0, 1\}$. (We note that Lemma II.7 is in fact false for $\tau = 0$, with the parity function serving as a counterexample.) We do this by showing that the set of (ε, τ) for which $\operatorname{conv} \mathcal{C}_{\varepsilon,\tau}$ does *not* support reliable computation is closed.

Our proof that the "non-reliable computation region" is closed uses a characterization—which may be of independent interest—of those circuit models $\mathcal{C}$ whose convex hulls $\operatorname{conv} \mathcal{C}$ support reliable computation. More precisely, we formalize the relationship between amplification and classical fault-tolerant computation. We discuss this formalization more in the next section.

E. Equivalence between reliable computation and amplification

We say that a function $f : \{0, 1\}^n \to \{0, 1\}$ is an *amplifier* if it amplifies the probability of a 1 (resp. 0) when given as input i.i.d. bits which are slightly biased towards 1 (resp. 0). The relationship between amplifiers and reliable computation has been implicitly exploited in previous work. However, making this relationship explicit (which turns out to be somewhat involved), is helpful in proving Theorem II.5. Moreover, it has applications to nonlocal games, as discussed in Sections II-B and II-C. We hope that this formalization will be useful for other questions in fault-tolerant computation.

We establish the relationship between reliable computation and amplification with the following theorem. Note that this theorem applies to arbitrary circuits, not just formulas.

Theorem II.8 (Equivalence between reliable computation and amplification)**.** *Let $\mathcal{C}$ denote a circuit model. Then $\operatorname{conv} \mathcal{C}$ supports reliable computation if and only if $\operatorname{conv} \mathcal{C}$ contains both an amplifier and a $\neg_\kappa$ gate for $\kappa < 1/2$.*

Further, given a circuit model $\mathcal{C}$ such that $\operatorname{conv} \mathcal{C}$ supports reliable computation, there exists a constant s such that for any function $f : \{0, 1\}^n \to \{0, 1\}$ computable by a depth-d circuit of noiseless NAND gates, f can be computed by a depth-$(s \cdot d)$ circuit in $\operatorname{conv} \mathcal{C}$ with failure probability bounded away from $1/2$.

Theorem II.8 may be viewed as a generalization of a number of results of a similar flavor which have been proven, explicitly or implicitly, for specific circuit models [14], [15], [16], [17], [17], [18], [19], [5]. However, to the best of our knowledge no equivalence in this generality has been stated before; perhaps this is because previous work has not explicitly considered convex hulls of circuit models.

The depth statement in Theorem II.8 (the "Further" clause) is closely related to many classical results on the depth and size complexity overhead for fault-tolerance, for example [20], [21]. Our result in Theorem II.8 differs from previous work in that it holds whenever reliable computation is possible (as opposed to for some fixed noise level or gate set). To the best of our knowledge, the depth statement in Theorem II.8 is not an immediate consequence of prior work.

One direction of Theorem II.8 is straightforward. Supposing $\operatorname{conv} \mathcal{C}$ supports reliable computation, it can reliably compute the majority function on k variables $\mathrm{Maj}^{(k)}$; for sufficiently large k this leads

to an amplifier in $\operatorname{conv}\mathcal{C}$.

The other direction is more involved. If $\operatorname{conv}\mathcal{C}$ contains an amplifier and a $\neg_\kappa$ gate for $\kappa < 1/2$, then one can construct a map that behaves similarly to a NAND gate, and use this map to reliably compute an arbitrary Boolean function. Since the depth of this approximate NAND gate is constant (for fixed κ), this means that the depth of any fault-tolerant formula is only a constant factor larger than the depth of the noise-free formula. We prove Theorem II.8 in [9]. For our purposes, Theorem II.8 is useful for two reasons. First, it is useful in formalizing the connection to nonlocal games, as discussed in Section II-B. Second, Theorem II.8 makes it easy to prove the following Theorem II.9, which provides the last part of the proof of Theorem II.5 that we outlined above in Section II-D.

Theorem II.9. *Let $\mathcal{C}_\varepsilon$ denote a circuit model on a gate set $\mathcal{G}$ which includes a noisy gate g_ε. Let $I \subseteq [0,1]$ denote the set of ε for which $\operatorname{conv}\mathcal{C}_\varepsilon$ does not support reliable computation (varying the noise on g_ε and keeping all other gates in $\mathcal{G}$ fixed). Then I is closed.*

Notice that Theorem II.9 directly implies that there exists a nonzero noise threshold for any circuit model that is closed under convex combinations and based on a functionally complete set of logic gates. The proof of Theorem II.9 is a simple consequence of Theorem II.8. Suppose that $\mathcal{C}_\varepsilon = \operatorname{conv}\mathcal{C}_\varepsilon$ supports reliable computation. Then Theorem II.8 implies there is an amplifier and a $\neg_\kappa$ gate for $\kappa < 1/2$ in $\mathcal{C}_\varepsilon$. Using some elementary analytical lemmas, one can show that these finite circuits retain their nature despite a sufficiently small "nudge" in the noise rate ε, and hence applying Theorem II.8 again, the resulting circuit model still supports reliable computation. We prove Theorem II.9 in [9].

III. Related Work

A. Axiomatization of Quantum Mechanics

Given its implications for the definition and behavior of information, there have been several proposals for sets of information-theoretic axioms that can be used to derive quantum mechanics. Examples include those of Hardy [22] as well as those of Mueller and Masanes [23]. While those efforts are enlightening in many ways, they don't directly address quantum nonlocality. However, the concise and uncontroversial requirement that "communication complexity is not trivial" is known to place stringent constraints on that nonlocality, so a line of work has considered whether the requirement could function as an axiom precisely delineating the limits of quantum mechanics [3], [5], [7].

The work of van Dam [3] established that the ability to win the CHSH game with probability 1 (that is, access to a so-called *Popescu-Rohrlich* (PR) box) causes communication complexity to be trivial. As discussed above, the work of Brassard, Buhrman, Linden, Méthot, Tapp, and Unger [5] extended this result to apply to success probability greater than 0.908. However, there is still a gap between this value and $\omega_Q(CHSH) \approx 0.8536$.

The work of Forster, Winkler and Wolf [6] shows that certain superquantum correlations on the boundary of the NS polytope can be distilled into perfect PR boxes, and thus also collapse communication complexity.[6] Brunner and Skrzypczyk [7] considered adding noise to these correlations and extended this set of superquantum correlations that collapse communication complexity (which we will call the *distillable set*). Allcock, Brunner, Linden, Popescu, Skrzypczyk, and Vértesi [10] introduced the notion of a set of correlations remaining closed under wirings; they exhibited convex sets of correlations without this property.

The distillable set has points arbitrarily close to a vertex of the polytope of classical correlations C. Unfortunately this does not produce a nonlocal game whose quantum value is exactly limited by the requirement that communication complexity be nontrivial. Geometrically, this is because supporting hyperplanes of Q at points of intersection between Q and the boundary of the distillable set are supporting hyperplanes of NS itself. Therefore

[6] Distillation protocols are descriptions of wirings that combine multiple "weaker" nonlocal correlations such that the new correlation is more useful (e.g., for playing the CHSH game).

superquantum advantage at such games is not possible in any (possibly superquantum) theory.

Other works have computed the optimality of distillation protocols, shown impossibility results within restricted settings (e.g., for nonadaptive procedures), and exhibited closed sets of superquantum correlations [24], [25], [26], [8], [27].[7] One might have hoped to improve the construction of Brassard *et al.* [5] by first distilling slightly superquantum noisy PR boxes to obtain better ones, which would then collapse communication complexity. However, prior work has yet to discover a distillation procedure for noisy PR boxes, or to rigorously rule out that one exists.

Navascués, Guryanova, Hoban, and Acín [28] introduced the set $\tilde{Q}$ of "almost-quantum" correlations, which strictly contains Q and has nontrivial communication complexity. This result implies there are many superquantum correlations which do not collapse communication complexity, and that other principles beyond "communication complexity is nontrivial" are required to discriminate points between Q and $NS \setminus Q$. On the other hand, since $\omega_{\tilde{Q}}(CHSH) = \omega_Q(CHSH)$, this left open the possibility that $\omega_Q(CHSH)$ is the maximum value consistent with nontrivial communication complexity.

B. Fault-tolerant Computation from Noisy Gates

Fault-tolerant computation by circuits has been studied extensively since von Neumann's work in the 1950's. A central question is how noisy the gates can get before reliable computation is impossible. In general, stronger bounds on the noise threshold have been obtained for formulas rather than general circuits; Theorem II.5 holds only for formulas, although we conjecture that a similar result holds for circuits as well. Almost all work[8] that we are aware of in fault-tolerant computation focuses on symmetric noise.

Modern work in the symmetric case goes back to the work of von Neumann in 1956, who showed that reliable computation is possible using noisy 3-majority gates which fail independently with probability $\varepsilon \leq 0.0073$ [14]. Since then, there has been a great deal of work; we summarize the best results in this setting in Table I.

To gain some intuition for these results, it is helpful to understand amplification, which we discussed in Sections I and II and which we define formally in the complete version of this paper [9]. All of the positive results that we are aware of go through amplifiers. That is, these works construct an amplifier out of the target gate set and then use that, perhaps along with other gates, to establish a method for reliable computation. For example, von Neumann [14] used a noisy 3-input majority gate $\mathrm{Maj}_\varepsilon^{(3)}$ as an amplifier; both Hajek and Weller [16] as well as Evans and Schulman [17] also used $\mathrm{Maj}_\varepsilon^{(k)}$ as an amplifier, and used noisy $\mathrm{XNAND}_\varepsilon$ gates along with this amplifier to improve von Neumann's result to give a sharp threshold for k-input gates for odd k. Evans and Pippenger [18] and Unger [19] used the amplifier

$$\mathrm{NAND}_\varepsilon(\mathrm{NAND}_\varepsilon(X_0, X_1), \mathrm{NAND}_\varepsilon(X_2, X_3)), \tag{2}$$

along with more $\mathrm{NAND}_\varepsilon$ gates to establish reliable computation for any $\varepsilon < \varepsilon_0 = \frac{3-\sqrt{7}}{4}$. The work of [18] showed a matching upper bound for reliable computation by formulas of noisy $\mathrm{NAND}_\varepsilon$ gates, assuming noisy inputs, showing that reliable computation is impossible when $\varepsilon > \varepsilon_0$ under these assumptions. Finally [19] extended the impossibility result to also include the case where $\varepsilon = \varepsilon_0$, removed the assumption that the inputs are noisy, and generalized the result to computation by formulas of all 2-input ε-noisy gates. As we explain further in Section III-A, the limit on nonlocality from nontrivial communication complexity is derived in [5] using the following amplifier:

$$((X_0 \oplus X_2) \wedge_\varepsilon (X_0 \oplus X_1)) \oplus X_0 \tag{3}$$

Because all of the positive results go through amplifiers, it is natural to wonder whether there is a deeper connection between the amplifiers and reliable computation in a circuit model, and this is what we show in Theorem II.8. Although such a

[7] [8] also extended the distillable set.

[8] We note that one exception to the symmetric noise paradigm is [29] which shows that if an adversary gets to decrease the noise heterogeneously from gate to gate, fault tolerant computation actually becomes harder.

Noise Model	Source	Circuit model	Bounds on threshold ε_0
Symmetric Noise	[15]	All circuits of ε-noisy gates of fan-in k	$\varepsilon_0 \leq \frac{1}{2} - \frac{1}{2\sqrt{k}}$
	[16], [17]	Formulas of ε-noisy gates of odd fan-in k	$\varepsilon_0 = \frac{1}{2} - \frac{2^{k-1}}{k\binom{k-1}{k/2-1/2}}$
	[18], [19]	Formulas of ε-noisy gates of fan-in 2	$\varepsilon_0 = \frac{3-\sqrt{7}}{4} \approx 0.08856$
Asymmetric Noise	[5]	Formulas of $\{\wedge_\varepsilon, \oplus_0\}$ gates	$\varepsilon_0 \geq 1/6$
	This work	Formulas of $\{\wedge_\varepsilon, \oplus_0\}$ gates	$\varepsilon_0 = 1/6$

Table I: Summary of best results on thresholds in both the symmetric and asymmetric case. Above, ε_0 represents the noise threshold so that if $\varepsilon < \varepsilon_0$ then reliable computation is possible, but if $\varepsilon \geq \varepsilon_0$ then it is impossible.

connection is implicit in prior work, to the best of our knowledge it has not been made rigorous. This may be because the equivalence is easier to prove using $\mathrm{conv}\,\mathcal{C}$ rather than $\mathcal{C}$ itself.

IV. Conclusion and Future Work

We investigated the extent to which the axiom "communication complexity is nontrivial" can explain the quantum value of nonlocal games, along the way developing new results about reliable classical computation with noisy gates. On the quantum side, we have shown that there is a game G so that $\omega_Q(G)$ is precisely explained by the axiom "communication complexity is nontrivial"; and we have ruled out the approach of [5] to show a similar statement for the CHSH game. On the reliable computation side, we have shown that the class $\mathcal{C}_\varepsilon$ of formulas made from $\wedge_\varepsilon$ and $\oplus_0$ gates does not support reliable computation for any $\varepsilon \in [1/6, 5/6]$. Combined with previous work of [5], this shows that the noise threshold for $\mathcal{C}_\varepsilon$ is exactly $1/6$. To prove our results, we have developed new tools for reasoning about fault-tolerant computation with asymmetric noise, including formalizing the tight relationship between amplifiers and fault-tolerant computation.

We conclude with a few open questions and directions for future work.

1) **Establishing that "communication complexity is nontrivial" is *not* enough to explain $\omega_Q(CHSH)$.** We have shown that the approach of Brassard *et al.* in [5] cannot be pushed further, in the sense that there is no amplifier in $\mathcal{C}_\varepsilon$ better than the one they have found. However, this does not rule out all approaches; in particular, it could be that there is a way to use the CHSH correlation in way other than to create noisy AND gates. It would be interesting to rule out any approach (or to find an approach that works!)

2) **An analogous result for circuits.** As with previous results about formulas (eg, [18], [19]), we conjecture that the same threshold of $\varepsilon = 1/6$ that we have proved for formulas also holds for circuits. The assumption of formulas only comes into the proof of Lemma II.7, where we use the fact that the noise in the subtrees beneath two different inputs is independent. It would be interesting to see if this assumption could be relaxed by investigating the nature of the dependencies which arise in general circuits.

3) **Results for general asymmetric gate noise.** We have studied the gate set $\{\wedge_\varepsilon, \oplus_\tau\}$ for the case that $\tau = 0$. However, it remains open for $\tau > 0$. The parameter regime where $(\varepsilon, \tau) \in \left(\frac{3-\sqrt{7}}{4}, 1/6\right) \times \left(0, \frac{3-\sqrt{7}}{4}\right)$ is of particular interest. Indeed, we understand what happens on the boundaries of this region, but do not know what happens in the interior.

4) **Relationship to quantum fault-tolerant computation.** Theorem II.8 may be viewed as an upper bound on the overhead required for reliable computation: regardless of the noise rate, there will only ever be a constant blow-up in the

depth of the circuit when using noisy gates to compute reliably. It is interesting to consider the analogous question in quantum computation, where realistic gate implementations will have significant gate noise necessitating fault-tolerance techniques in order to scale. Despite the resulting enormous amount of work on fault-tolerant quantum computation, it is not known whether a corresponding statement about constant blow-up in depth applies in the quantum setting. It is possible that there are multiple distinct thresholds in the quantum case: one noise threshold below which quantum circuits can reliably compute with minimal overhead, and a higher noise threshold below which quantum circuits can reliably compute at all. Indeed, a trivial version of the statement is almost certainly true; one limit of maximally asymmetric gate noise simply turns a quantum computer into a noiseless classical computer. Such a computer could simulate quantum computation but only with exponential overhead as far as we know. More interestingly, the possibility of multiple thresholds is supported, for example, by the work of [30] which shows that circuits of sufficiently noisy quantum gates are efficiently simulatable by a classical circuit. It is also consistent with the best known constructions for fault-tolerant quantum computing.[9] Specifically, in order to obtain fault-tolerance with a constant factor overhead in quantum circuit depth, the best current construction has a threshold that is orders of magnitude worse than thresholds from proposals with super-constant overhead [31], [32] .

It may well be that a quantum version of Theorem II.8 exists, meaning there remain major improvements to be found in quantum fault tolerance that will achieve constant depth overhead at high noise rates. That would be an exciting and likely technologically important discovery. On the other hand, the story may simply be more complicated in the quantum setting, with multiple thresholds depending on the scaling of the overhead cost, which would be a sharp contrast to what happens for reliable classical computation.

[9]The usual model allows noiseless classical computation on the side to perform syndrome calculations for error correction.

ACKNOWLEDGMENT

We thank Li-Yang Tan for helpful discussions. We thank the Stanford Research Computing Center and Google for providing computing resources. We also thank anonymous reviewers for helpful comments and suggestions.

REFERENCES

[1] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Physical review letters*, vol. 23, no. 15, p. 880, 1969.

[2] B. S. Cirel'son, "Quantum generalizations of bell's inequality," *Letters in Mathematical Physics*, vol. 4, no. 2, pp. 93–100, 1980.

[3] W. van Dam, "Implausible consequences of superstrong nonlocality," *Natural Computing*, vol. 12, no. 1, pp. 9–12, 2013.

[4] R. Cleve, personal communication.

[5] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, "Limit on nonlocality in any world in which communication complexity is not trivial," *Physical Review Letters*, vol. 96, no. 25, p. 250401, 2006.

[6] M. Forster, S. Winkler, and S. Wolf, "Distilling nonlocality," *Physical review letters*, vol. 102, no. 12, p. 120401, 2009.

[7] N. Brunner and P. Skrzypczyk, "Nonlocality distillation and postquantum theories with trivial communication complexity," *Physical review letters*, vol. 102, no. 16, p. 160403, 2009.

[8] P. Høyer and J. Rashid, "Optimal protocols for nonlocality distillation," *Physical Review A*, vol. 82, no. 4, p. 042118, 2010.

[9] N. Shutty, M. Wootters, and P. Hayden, "Tight limits on nonlocality from nontrivial communication complexity; a.k.a. reliable computation with asymmetric gate noise," 2018. [Online]. Available: https://arxiv.org/abs/1809.09748

[10] J. Allcock, N. Brunner, N. Linden, S. Popescu, P. Skrzypczyk, and T. Vértesi, "Closed sets of nonlocal correlations," *Physical Review A*, vol. 80, no. 6, p. 062107, 2009.

[11] N. D. Mermin, "Simple unified form for the major no-hidden-variables theorems," *Phys. Rev. Lett.*, vol. 65, pp. 3373–3376, Dec 1990. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.65.3373

[12] A. Peres, "Incompatible results of quantum measurements," *Physics Letters A*, vol. 151, no. 3, pp. 107 – 108, 1990. [Online]. Available: http://www.sciencedirect.com/science/article/pii/037596019090172K

[13] N. Pippenger, "Reliable computation by formulas in the presence of noise," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 194–197, 1988.

[14] J. von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," *Automata studies*, vol. 34, pp. 43–98, 1956.

[15] W. S. Evans and L. J. Schulman, "Signal propagation and noisy circuits," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2367–2373, 1999.

[16] B. Hajek and T. Weller, "On the maximum tolerable noise for reliable computation by formulas," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 388–391, 1991.

[17] W. S. Evans and L. J. Schulman, "On the maximum tolerable noise of k-input gates for reliable computation by formulas," *IEEE Transactions on Information Theory*, vol. 49, no. 11, pp. 3094–3098, 2003.

[18] W. Evans and N. Pippenger, "On the maximum tolerable noise for reliable computation by formulas," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 1299–1305, 1998.

[19] F. Unger, "Noise threshold for universality of 2-input gates," in *IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2007, pp. 1901–1905.

[20] N. Pippenger, "On networks of noisy gates," in *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*. IEEE, 1985, pp. 30–38.

[21] R. L. Dobrushin and S. Ortyukov, "Upper bound on the redundancy of self-correcting arrangements of unreliable functional elements," *Problemy Peredachi Informatsii*, vol. 13, no. 3, pp. 56–76, 1977.

[22] L. Hardy, "Quantum theory from five reasonable axioms," *arXiv preprint quant-ph/0101012*, 2001.

[23] M. P. Mueller and L. Masanes, "Information-theoretic postulates for quantum theory," in *Quantum Theory: Informational Foundations and Foils*. Springer, 2016, pp. 139–170.

[24] D. D. Dukaric and S. Wolf, "A limit on non-locality distillation," *arXiv preprint arXiv:0808.3317*, 2008.

[25] A. J. Short, "No deterministic purification for two copies of a noisy entangled state," *Phys. Rev. Lett.*, vol. 102, p. 180502, May 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.102.180502

[26] M. Forster, "Bounds for nonlocality distillation protocols," *Phys. Rev. A*, vol. 83, p. 062114, Jun 2011. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.83.062114

[27] B. Lang, T. Vértesi, and M. Navascués, "Closed sets of correlations: answers from the zoo," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 42, p. 424029, 2014.

[28] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, "Almost quantum correlations," *Nature communications*, vol. 6, p. 6288, 2015.

[29] F. Unger, "Better gates can make fault-tolerant computation impossible," in *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 17, 2010, p. 164.

[30] S. Virmani, S. F. Huelga, and M. B. Plenio, "Classical simulability, entanglement breaking, and quantum computation thresholds," *Physical Review A*, vol. 71, no. 4, p. 042328, 2005.

[31] O. Fawzi, A. Grospellier, and A. Leverrier, "Constant overhead quantum fault-tolerance with quantum expander codes," in *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2018, pp. 743–754.

[32] D. Gottesman, "Fault-tolerant quantum computation with constant overhead," *Quantum Information & Computation*, vol. 14, no. 15-16, pp. 1338–1372, 2014. [Online]. Available: http://www.rintonpress.com/xxqic14/qic-14-1516/1338-1371.pdf