New Limits on Fault-Tolerant Quantum Computation

Harry Buhrman*
CWI and
U of Amsterdam
buhrman@cwi.nl

Richard Cleve†
U of Waterloo and
Perimeter Institute
cleve@cs.uwaterloo.ca

Monique Laurent ‡
CWI, Amsterdam
M.Laurent@cwi.nl

Noah Linden§
U of Bristol
n.linden@bristol.ac.uk

Alexander Schrijver
CWI and
U of Amsterdam
Lex.Schrijver@cwi.nl

Falk Unger*
CWI, Amsterdam
Falk.Unger@cwi.nl

Abstract

We show that quantum circuits cannot be made fault-tolerant against a depolarizing noise level of $\hat{\theta} = (6 - 2\sqrt{2})/7 \approx 45\%$*, thereby improving on a previous bound of* 50% *(due to Razborov [18]). More precisely, the circuit model for which we prove this bound contains* perfect *gates from the Clifford group (CNOT, Hadamard,* S*,* X*,* Y*,* Z*) and arbitrary additional one-qubit gates that are subject to depolarizing noise* $\hat{\theta}$*. We prove that this set of gates cannot be universal for arbitrary (even classical) computation, from which the upper bound on the noise threshold for fault-tolerant quantum computation follows.*

1 Introduction

In the past decade, quantum computing has attracted much attention because of its ability to efficiently solve problems for which no efficient classical algorithms are known. Significant research efforts are dedicated to physically realizing quantum computers. A fundamental problem is to cope with noise, which creates major difficulties in storing and operating on quantum states reliably. A key advance was the realization that quantum error correcting codes [21, 24] exist and fault-tolerant quantum computation [22] is possible for a number of reasonable error models. Subsequent results have improved on the first fault-tolerant schemes, proving better and better bounds on the noise tolerable in quantum computation (e.g. [6, 3]). Recent results suggest that fault-tolerant quantum computation is possible with gates that have as much as 3% of depolarizing errors [12], but there is no rigorous proof so far.

In this paper we will concentrate on the opposite task of proving that, for certain noise levels, quantum computation is *impossible*. Our main result is as follows: Let CLIFFORD be the set of all (noiseless) Clifford gates

$$\text{CNOT}_2^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{1}$$

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}.$$

The Gottesman-Knill Theorem says that this set of gates can be efficiently simulated classically (see also [1]), so they are probably not universal for quantum computation. On the other hand, it is known that CLIFFORD together with *any* other one-qubit gate, not generated by the gates in CLIFFORD, form a universal set of gates for quantum computation [23, 14]. We show however, that such additional one-qubit gates should not be too noisy. More precisely, let CLIFFORD* be CLIFFORD augmented with arbitrary one-qubit gates with depolarizing error at least $\hat{\theta} = (6 - 2\sqrt{2})/7 \approx 45\%$. Then this set of gates is no longer capable of computing arbitrary functions and thus is not universal. In other words, fault-tolerant quantum computation cannot be performed if there is this level of noise. Additionally we show that, among all one-qubit gates that augment CLIFFORD, the so-called $\pi/8$-gate (see end of Section 3) is the type of gate that requires the most noise to render it incapable of universal quantum computation by our approach. That is, for other augmenting gates (e.g., $\pi/16$-gates), our approach will yield stronger bounds on the tolerable level. Our results also yield a simple proof that not

*Supported by the EU project RESQ IST-2001-37559 and the NWO vici project 2004-2009.
†Supported in part by Canada's NSERC, CIAR, MITACS, and the US ARO.
‡Supported by the Netherlands Organization for Scientific Research NWO 639.032.203.
§Supported by the EU project RESQ IST-2001-37559 and UK EPSRC IRC in Quantum Information Processing.

all classical functions can be computed using Clifford gates (complementing results in [1]). In particular, in Corollary 1, we show that a boolean function which can be computed by Clifford circuits can be written as the parity of a subset of input bits.

The main idea of our approach is as follows. Assume we have a Clifford circuit C with n classical input bits $x = x_1, \ldots, x_n$ and one dedicated output qubit that, when measured in the computational basis, yields the output of the computation of C on x. Suppose now that the input is partitioned over two parties, Alice and Bob, such that Alice has k bits of x and Bob has $n - k$ bits. We first show how Alice, with the help of Bob, can compute the value of C on x with just a single classical bit of communication (Lemma 1). From this it follows that Clifford circuits can at the very best compute only those functions that require for any partition of the inputs a single bit of communication, and it is well known that many functions require more than one bit of communication. Next, we show in Lemma 2 how probabilistic mixtures of Clifford gates can be used to simulate any single qubit unitary gate, that has noise $\hat{\theta}(\approx 45\%)$. The proof of our Lemma relies on solving an optimization problem related to the Clifford polytope, defined as the convex hull of the set $\mathcal{C} \subseteq \mathbb{R}^{3\times 3}$ of Clifford rotation matrices in $\mathbb{R}^3$. Here, the matrices $\mathcal{C}$ are essentially the one-qubit Clifford gates in Bloch sphere representation.

Combining Lemmas 1 and 2, we get that all circuits with CLIFFORD*-gates and with respect to any distribution of the inputs can be computed by Alice and Bob with a single bit of communication (Lemma 3). Using the fact that there are functions which require communication more than one bit, we get our main result (Theorem 1): The set of gates in CLIFFORD* cannot be universal. We also generalize our result to the case that the inputs are quantum states.

The idea that a noisy 1-qubit gate can be simulated by a probabilistic mixture of Clifford appeared first in Virmani *et al.* [26]. The approach we take here though is an extension to quantum fault tolerant computation of the work by Brassard et al [4], where they exhibit an upper bound on the noise threshold for classical fault tolerant computation, using lower bounds on quantum communication complexity and the non-local CHSH correlation.

We want to point out that section 3 can be read independently of the preceding section. It shows that gates from CLIFFORD*, together with all stabilizer operations and classical co-processing are classically simulatable and thus probably not quantum-universal.

1.1 Related Work

There are only a few other results concerning the limits of fault-tolerant computation. These are not all strictly comparable to each other and our result; nevertheless, we review them and make some comparisons. See the introduction of [18] for some remarks that motivate the analysis of thresholds for fault-tolerant quantum computation.

The first results on upper bounds of the threshold decoherence rate were obtained by showing that quantum computers with faulty gates can be simulated efficiently on a classical computer. The first to prove one of these results were Aharonov and Ben-Or [2], with the value 97% for the noise. Later Harrow and Nielsen [11] showed that if 74% of depolarizing noise is applied to each output qubit of each gate, then (faulty) two-qubit gates cannot produce entanglement. They concluded that circuits containing only one- and two-qubit gates with depolarizing noise at least 74% can be simulated efficiently on a classical computer.

An improvement of this is due to Virmani *et al.* [26] who show that the set consisting of CNOT with depolarizing noise at least 67% and arbitrary 1-qubit gates is efficiently simulatable classically. In this paper they also introduce the interesting idea that sufficiently noisy 1-qubit gates can be simulated by Clifford gates; we build on and extend this idea in this paper. We note however, that their strongest results are for a restricted class of gates (ones which are diagonal in the computational basis) and dephasing or worst-case noise. They prove that $(\sqrt{2}-1)/\sqrt{2} \approx 29\%$ dephasing noise is enough to make these diagonal gates a mixture of Clifford operations[1]. We extend their results by considering all 1-qubit gates. Note also that dephasing noise is only symmetric around the z-axis, which is natural when considering diagonal gates. Our noise bounds are with respect to depolarizing noise, which is symmetric in all directions, and hence appropriate when considering arbitrary one-qubit unitaries.

Note that all these results do not exclude the possibility that quantum circuits with high noise can still do universal classical computations; our results imply this.

The only prior result of this latter type is due to Razborov [18], where a 50% upper bound on the noise threshold for depolarizing noise on qubits for circuits with two-qubit gates is obtained (and a weaker bound for k-qubit gates). The argument in [18] is essentially that, at this noise level, any superlogarithmic-depth quantum circuit (with constant error rate per qubit per time step) will be overwhelmed by the noise and produce a statistically meaningless outcome. Thus, under the complexity theoretic assumption $BQP \neq QNC^1$, there are sets in BQP which can be computed with this noise level. We note that it is shown in [8] that in fact log-depth quantum circuits can perform interesting feats, including efficient integer factorization (if combined by classical polynomial-time pre- and post-processing). Our error model is in most respects weaker than that of [18] (since our qubit errors are only occurring at the completion of non-CLIFFORD gates) and our bound of

[1] They define dephasing noise as $\rho \mapsto 1/2(\rho + Z\rho Z)$.

$\approx 45\%$ is below 50%. In fairness, there is a sense in which the bound in [18] is stronger: it permits arbitrary (noisy) two-qubit gates; whereas, our only two-qubit gates are (perfect) CNOT gates.

Finally, we note that our work is related to, and partly stimulated by, the circle of ideas surrounding measurement-based quantum computation that was largely initiated by [10, 17].

2 Preliminaries and Notation

E_{ij} is the all-zero matrix, except for the entry i, j which is equal to 1. We also write $+$ for $+1$ and $-$ for -1. For matrices $A, B \in \mathbb{R}^{3\times 3}$ we define the inner product $\langle A, B\rangle$ as:

$$\langle A, B\rangle = tr(A^T B) = \sum_{i,j\in\{1,2,3\}} a_{ij} b_{ij}.$$

The following fact is used repeatedly: $\langle A, BC\rangle = \langle B^T A, C\rangle$ for $A, B, C \in \mathbb{R}^{3\times 3}$.

By A^* we denote the conjugate transpose of matrix A.

An *n-qubit state* (or density matrix) ρ is a matrix $\rho \in \mathbb{C}^{2^n\times 2^n}$ with the properties $tr(\rho) = 1$, $\rho = \rho^*$ (Hermiticity) and ρ is positive semi-definite. An *n-qubit operation* (or *gate*) is a unitary matrix $U \in \mathbb{C}^{2^n\times 2^n}$, i.e., $U^*U = \mathbb{I}$. For such n-qubit state ρ and n-qubit operation U the application of U to ρ results in the state $U\rho U^*$.

2.1 Bloch-vector representation

In our further analysis it will be convenient to use the Bloch-sphere representation of 1-qubit states and 1-qubit operations, which we review now (see e.g. Section 4.2 and Chapter 8 in [15]).

For $\mathbf{r} \in \mathbb{R}^3$ define $\mathbf{r}\cdot\sigma = r_x X + r_y Y + r_z Z$, where $\sigma = (X, Y, Z)$ is the vector of Pauli matrices

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \; Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then, all 1-qubit density matrices ρ can be uniquely written in the form

$$\rho = \frac{\mathbb{I} + \mathbf{r}\cdot\sigma}{2} = \frac{\mathbb{I} + r_x X + r_y Y + r_z Z}{2},$$

where $\mathbf{r} \in \mathbb{R}^3$ and $||\mathbf{r}|| = \sqrt{r_x^2 + r_y^2 + r_z^2} \le 1$. We call $\mathbf{r}$ the *Bloch vector* of ρ.

For $\mathbf{n} \in \mathbb{R}^3$ with $||\mathbf{n}|| = 1$ and $\theta \in \mathbb{R}$ we define

$$U_{\mathbf{n}}(\theta) = \exp(-i\theta\mathbf{n}\cdot\sigma/2) = \cos(\theta/2)\mathbb{I} - i\sin(\theta/2)\mathbf{n}\cdot\sigma.$$

We first note that $U_{\mathbf{n}}(\theta)U_{\mathbf{n}}(\theta)^* = \mathbb{I}$, i.e., $U_{\mathbf{n}}(\theta)$ is unitary. Second, let the result of the quantum operation $U_{\mathbf{n}}(\theta)$ applied to state $\rho = \mathbb{I}/2 + \mathbf{r}\cdot\sigma/2$ be $\rho' = U_{\mathbf{n}}(\theta)^*\rho U_{\mathbf{n}}(\theta) = \mathbb{I}/2 + \mathbf{r}'\cdot\sigma/2$. Then $\mathbf{r}'$ is the image of rotating $\mathbf{r}$ around $\mathbf{n}$ by an angle θ. Third, all 1-qubit unitaries U can be written as

$$U = U_{\mathbf{n}}(\theta)$$

with $\mathbf{n} \in \mathbb{R}^3, \theta \in \mathbb{R}$ and $||\mathbf{n}|| = 1$ (ignoring an unimportant phase factor $\alpha \in \mathbf{C}$ with $|\alpha| = 1$).

Thus, one-qubit states and unitaries are isomorphic to vectors ,resp., rotations in $\mathbb{R}^3$. The set of all rotations in $\mathbb{R}^3$ is the group $SO(3)$.[2] We introduce some notation reflecting this isomorphism. For unitary $U \in \mathbb{C}^{2\times 2}$ we let $R_U \in SO(3)$ be the corresponding rotation matrix. We get a reverse operation (up to phase factors) by fixing one mapping $f : SO(3) \to \mathbb{C}^{2\times 2}$ with the property that for all unitary $U \in \mathbb{C}^{2\times 2}$ it holds $f(R_U) = \alpha U$ for some $\alpha \in \mathbb{C}$, $|\alpha| = 1$. We then write $U_R = f(R)$.

This can be extended to probabilistic mixtures of quantum operations. Let $\{p_i\}$ be a probability distribution, i.e., $\sum_i p_i = 1$ and $0 \le p_i$, and let $U_i \in \mathbb{C}^{2\times 2}$ be a 1-qubit unitary with corresponding Bloch representation $R_i \in \mathbb{R}^{3\times 3}$. Then the quantum operation E in which each U_i is applied with probability p_i has Bloch-representation $R_E = \sum_i p_i R_i$.

2.2 Noise

There are several models of noise considered in the literature. The most common one, which we consider too, is *depolarizing noise*. A 1-qubit state ρ to which depolarizing noise p is applied, becomes

$$(1-p)\rho + p\mathbb{I}/2.$$

Thus, with probability $1-p$ the state is not changed and with probability p the state is replaced with the completely mixed state.

It is not hard to see that applying depolarizing noise p to $\rho = \mathbb{I}/2 + \mathbf{r}\cdot\sigma/2$ yields $\rho' = \mathbb{I}/2 + \mathbf{r}'\cdot\sigma/2$, with $\mathbf{r}' = (1-p)\mathbf{r}$. So, this noise shrinks the Bloch vector of a state to $(1-p)$ of its original length.

We say that a 1-qubit gate implements the unitary U with noise p if it transforms states ρ into

$$(1-p)U\rho U^* + p\mathbb{I}/2. \tag{2}$$

This quantum operation can be seen as a two-stage process, in which first U and then depolarizing noise is applied. Let $R_U \in \mathbb{R}^{3\times 3}$ be the rotation matrix corresponding to the unitary U. Then this noisy quantum operation has Bloch-representation $(1-p)R_U$, i.e., it rotates a Bloch vector and scales it by a factor $1-p$.

For 1-qubit gates and depolarizing noise, the two representations are (up to unimportant global phase factors) equivalent. (See Section 8.3 in [15] for more details.)

[2] This group will play a prominent role in the proof of Lemma 2, where some more notation can be found.

2.3 Clifford group

The (n-qubit) Clifford group contains all unitary operations that can be written as a product of tensor products of S, H and $CNOT$ (see Eq. (1)). The Clifford group contains also all Pauli operators X, Y, Z. We let CLIFFORD be the set of all Clifford gates. Let CLIFFORD* be the set of gates consisting of CLIFFORD and arbitrary 1-qubit gates which have depolarizing noise at least $\hat{\theta} = (6 - 2\sqrt{2})/7$.

For a state with Bloch vector $\mathbf{r}$ we get:

$$S\left(\frac{1}{2}\mathbb{I} + \frac{r_x}{2}X + \frac{r_y}{2}Y + \frac{r_z}{2}Z\right)S^* = \frac{1}{2}\mathbb{I} - \frac{r_y}{2}X + \frac{r_x}{2}Y + \frac{r_z}{2}Z$$

Let R_S be the Bloch representation of S. Then R_S rotates Bloch vectors around the z-axis by $\pi/2$. In particular, the x-axis is mapped to $-y$ and y to x. For the Hadamard-gates we similarly have

$$H\left(\frac{1}{2}\mathbb{I} + \frac{r_x}{2}X + \frac{r_y}{2}Y + \frac{r_z}{2}Z\right)H^* = \frac{1}{2}\mathbb{I} + \frac{r_z}{2}X - \frac{r_y}{2}Y + \frac{r_x}{2}Z.$$

So the Bloch representation R_H of H negates the y-coordinate of a Bloch vector and swaps the x and z-coordinates, i.e., it is a rotation by π around the axis $1/\sqrt{2}(1, 0, 1)$.

We define $\mathcal{C}$ as the set of matrices which can be generated from R_S and R_H. A $C \in \mathcal{C}$ is called a *Clifford (rotation) matrix*. It is not hard to see that $\mathcal{C}$ contains exactly those rotations which map axes to axes (or their opposite). Those C have in each row and column exactly one non-zero entry, which must be either $+1$ or -1, and $det(C) = 1$. Note that $\mathcal{C}$, being isomorphic to the 1-qubit Clifford group, is a group under matrix multiplication. Examples of Clifford matrices are

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

2.4 Communication Complexity

The setting for this is the following: Assume two separated parties, Alice and Bob, where Alice is given $x \in \{0,1\}^{m_A}$ and Bob $y \in \{0,1\}^{m_b}$, want to compute $f(x,y)$ for some fixed function $f : \{0,1\}^{m_A} \times \{0,1\}^{m_B} \to \{0,1\}$. We want that at least one party learns the result $f(x,y)$. In order to achieve this they can communicate bits, according to a predefined protocol. The *deterministic communication complexity* $C_f(n)$ of f is the smallest number c such that each protocol which always computes the correct result, needs at least c bits of communication for at least one input $x, y \in \{0,1\}^n$.

It is well-known that there are functions f where $C_f(n)$ is n, for example the inner product function (see [13]). The above can be extended to *randomized* communication, where the parties are additionally provided with a source which sends a sequence of random bits to Alice and the same sequence to Bob. The final result only has to be correct with some probability $1 - \epsilon$ for $\epsilon < 1/2$. The minimum number of bits needed to be communicated such that the output is correct with probability at least $1 - \epsilon$ is denoted by $C_f^\epsilon(n)$. However, also in this randomized setting there are "hard" functions. For example, it is known that the inner product function has randomized communication complexity $n - O(\log(1/\delta))$, if the outputs have to be correct with probability at least $1/2 - \delta$ (see also [13]).

For functions $f : \{0,1\}^n \to \{0,1\}$ which depend only on one input string and any $S \subseteq \{1, \ldots, n\}$ let $C_f(n, S)$ be the communication complexity of f if the bits with indices in S are given to Alice and all others to Bob. As in [13] we then define the *worst-case partition communication complexity* as $C_f(n) = \max_{S \subseteq \{1,\ldots,n\}} C_f(n, S)$. In [25] this is called symmetric communication complexity.

3 The power of Clifford circuits

We are now ready to explain the idea of simulating Clifford circuits.

Lemma 1. *Let $f : \{0,1\}^n \to \{0,1\}$ be a function that is computable with unbounded error*[3] *by a quantum circuit C that uses only gates from* CLIFFORD *, ancillas initialized to $|0\rangle$ and one single-qubit measurement in the computational basis, which determines the output. Then the deterministic communication complexity $C_f(n)$ is at most one bit.*

Proof. We begin by noting that each qubit can be represented by two *shares*: a *classical share* consisting of two bits, and a *quantum share* consisting of one qubit . When the classical share is ab and the quantum share $|\psi\rangle$, then the logical qubit that the shares encode is $X^a Z^b |\psi\rangle$.

Assuming that a set of qubits is encoded in this manner, the operations H, S, and CNOT can be applied to the logical qubits by separately performing operations on the shares that encode them (i.e., the logical qubits do not have to be reconstructed). The reason why this works is because for any Clifford operation $C = H, S, \text{CNOT}_2^1$ and any tensor product of Pauli operators P_1 there is a tensor product of Pauli operators P_2 with $CP_1 = P_2 C$. For example, to apply H to a logical qubit, the two bits that make up its classical share are swapped and H is applied to its quantum share. This works correctly because

$$\begin{aligned} HX^aZ^b|\psi\rangle &= HX^aHHZ^bHH|\psi\rangle \\ &= Z^aX^bH|\psi\rangle \\ &= (-1)^{a \wedge b} X^b Z^a H|\psi\rangle, \end{aligned} \tag{3}$$

[3] That means, that the output is only correct with probability strictly greater than $1/2$.

and $(-1)^{a\wedge b}$ is an irrelevant global phase.

To apply S to a logical qubit, the b-part of the classical share is updated to $b := a \oplus b$ and S is applied to its quantum share. This case can be verified by noting that

$$\begin{aligned} SX^aZ^b|\psi\rangle &= i^aX^aSZ^aZ^b|\psi\rangle \\ &= i^aX^aZ^{a\oplus b}S|\psi\rangle, \end{aligned} \tag{4}$$

where we note that i^a is a global phase.

To simulate the application of CNOT_2^1 gate[4] on two logical qubits, with classical shares a_1b_1 and a_2b_2, we update $a_2 := a_1 \oplus a_2$, $b_1 := b_1 \oplus b_2$ and CNOT_2^1 is applied to the two quantum shares. In this case, we omit the details but note that the correctness can be verified using the identities

$$\begin{aligned} \mathrm{CNOT}_2^1(X\otimes I) &= (X\otimes X)\mathrm{CNOT}_2^1 \\ \mathrm{CNOT}_2^1(I\otimes X) &= (I\otimes X)\mathrm{CNOT}_2^1 \\ \mathrm{CNOT}_2^1(Z\otimes I) &= (Z\otimes I)\mathrm{CNOT}_2^1 \\ \mathrm{CNOT}_2^1(I\otimes Z) &= (Z\otimes Z)\mathrm{CNOT}_2^1. \end{aligned} \tag{5}$$

We first describe a *probabilistic* communication protocol for f. Alice operates on the classical shares while Bob operates on the quantum shares. The initial shares are easy to construct: for each of Alice's input qubits $|x_j\rangle$, Alice sets her classical share to $a_j := x_j, b_j := 0$ and Bob sets his quantum share to $|0\rangle$; for each of Bob's input bits y_j, Alice sets her classical share to $a_j = b_j := 0$ and Bob sets his quantum share to $|y_j\rangle$. In this manner, Alice and Bob can simulate the execution of circuit C on input $|x\rangle|y\rangle|0\dots0\rangle$ *without any communication* to obtain the shares of the output qubits of C. For Bob to obtain the measured output qubit, Alice sends the first bit of her classical share, a_1, to Bob, who applies X^{a_1} to his quantum share and measures it (Alice need not send b_1, the second bit of the classical share, since Bob is performing a measurement in the computational basis).

Finally, to obtain a *deterministic* communication protocol for f, we note that Bob need not actually manipulate quantum information; rather, he can simulate his quantum registers and his operations with high enough precision on a classical computer. Then, upon receipt of the classical bit from Alice, he can exactly determine the output probabilities of his measurement to determine which outcome is more likely. □

The next Corollary characterizes exactly all functions computable by Clifford circuits. From Lemma 1 we get that this set is very limited and far from being universal.

Corollary 1. *All functions $f : \{0,1\}^n \to \{0,1\}$ which can be computed by a Clifford circuit, can be written in the form*

$$f(x_1\dots x_n) = c \oplus \bigoplus_{j\in S} x_j,$$

where $S \subseteq [n]$ is a subset of the input bits not depending on the input bits and $c \in \{0,1\}$.

Proof. It is clear that all functions f of this form can be computed by a Clifford circuit. We now also prove the reverse.

Let $f : \{0,1\}^n \to \{0,1\}$ be a function which can be computed by a Clifford circuit C. Then we can simulate this circuit as in Lemma 1, where we give Alice the whole input, i.e., $m_A = n$ and $m_B = 0$.

Inspecting the proof of Lemma 1 we see that in each step Alice always updates her a_i's and b_i's by computing the parity of two bits. So, the final bit she sends over, say a_i, is just the parity of some of the input bits. Thus we can write $a_i = \bigoplus_{j\in S} x_j$, for some $S \subseteq [n]$. Bob initializes all his quantum bits to $|0\rangle$, so he starts with the state $|\psi^0\rangle = |0\dots0\rangle$. Further, Bob just applies the circuit C to his state and measures the i-th qubit of $X^{a_i}C|\psi^0\rangle$ in the computational basis.

It is known that the probability for measuring 1 in a Clifford circuits is either 0, $1/2$ or 1 (see [15] page 463). It cannot be $1/2$ in our case, because that would mean that the circuit does not compute f. So, measuring the i-th bit of $C|\psi^0\rangle$ yields a bit $c \in \{0,1\}$ with certainty. But this means that $f(x) = c \oplus a_i = c \oplus \bigoplus_{j\in S} x_j$. □

We mention that Aaronson and Gottesman proved [1] that there is a log-space machine which transforms a Clifford circuit C into a classical circuit C' consisting only of CNOT and NOT gates, with the property that C accepts the all zero state $|0\rangle^{\otimes n}$ iff C' accepts the (classical) all zero input. Our corollary extends this slightly: For every Clifford circuit C computing a boolean function, there is an equivalent (for classical inputs) classical circuit which uses only NOT- and CNOT-gates. Using the result from [1] we see that we can compute the bit c in the proof of Corollary 1 in log-space and it is also clear that the circuit Alice uses to compute a_i can be computed in log-space.

Remark 1. It is straightforward to extend Lemma 1 to functions with m output bits, if the communication complexity of the function is also higher than m, resulting in a scheme that uses m bits of communication.

4 Simulating unitaries

We want to extend Lemma 1, by replacing CLIFFORD with CLIFFORD*. To do that we first show how one can simulate arbitrary 1-qubit gates with depolarizing noise $\hat\theta = (6-2\sqrt{2})/7$ with a probabilistic mixture of Clifford operations.

Lemma 2. *Let U be a 1-qubit unitary and E_U be the following noisy version of it*

$$\rho \mapsto E_U(\rho) = (1-\hat\theta)U\rho U^* + \hat\theta\mathbb{I}/2,$$

[4] control

for any $\rho \in \mathbb{C}^{2\times 2}$. Then there is a probability distribution $\{p_C\}$ over $\mathcal{C}$ such that for all $\rho \in \mathbb{C}^{2\times 2}$ we have

$$E_U(\rho) = \sum_{C\in\mathcal{C}} p_C U_C \rho U_C^*$$

and U_C is a Clifford operation corresponding to the Clifford rotation matrix C.

Proof. Using Section 2.1 the lemma can be reformulated equivalently in Bloch representation: For any $S \in SO(3)$ there is a probability distribution $\{p_C\}$ over $\mathcal{C}$ such that

$$(1-\hat{\theta})S = \sum_{C\in\mathcal{C}} p_C C.$$

We will prove this latter statement. Define the *Clifford polytope* as $P := \mathrm{conv}(\mathcal{C})$, i.e.

$$P = \left\{ S \mid S = \sum_{C\in\mathcal{C}} p_C C, p_C \geq 0, \sum_{C\in\mathcal{C}} p_C = 1 \right\} \quad (6)$$

as the convex hull of the 24 Clifford rotation matrices in $\mathbb{R}^{3\times 3}$. We have to prove

$$(1-\hat{\theta})S \in P \ \text{ for any } S \in SO(3). \quad (7)$$

For this we use the fact that the Clifford polytope can be alternatively described by its facet description:

$$P = \left\{ S \in \mathbb{R}^{3\times 3} \mid \langle F, S\rangle \leq 1 \ \text{ for all } F \in \mathcal{F} \right\}, \quad (8)$$

where

$$\mathcal{F} := \left\{ C_1 B C_2 | C_1, C_2 \in \mathcal{C}, B \in \{B_1, B_1^T, B_2\} \right\},$$
$$B_1 := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \ B_2 := \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

One can use the software from [9] for computing the facet description (8); we will give a proof in the paper version. In view of (8), our claim (7) is equivalent to

$$(1-\hat{\theta})\langle F, S\rangle \leq 1 \ \text{ for all } S \in SO(3), \ F \in \mathcal{F}. \quad (9)$$

Let $F \in \mathcal{F}$ of the form $F = C_1 B C_2$ where $C_1, C_2 \in \mathcal{C}$. As $\langle F, S\rangle = \langle C_1^T S C_2^T, B\rangle$ and $C_1^T S C_2^T \in SO(3)$, (9) is equivalent to

$$\langle S, B\rangle \leq \frac{1}{1-\hat{\theta}} = 2\sqrt{2} - 1 \quad (10)$$

for all $B \in \{B_1, B_2\}$, $S \in SO(3)$. The case $B = B_1$ is easy to handle: For $S \in SO(3)$, $\langle S, B_1\rangle = \sum_{i=1}^3 S_{i1} \leq \sqrt{3} < 2\sqrt{2} - 1$. We now show (10) for $B = B_2$. Write $S \in \mathbb{R}^{3\times 3}$ as

$$S = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

Well-known necessary and sufficient conditions for $S \in SO(3)$ are

$$\mathbf{a}^T\mathbf{b} = 0, \ \mathbf{c} = \mathbf{a}\times\mathbf{b}, \ \mathbf{a}^T\mathbf{a} = 1, \ \mathbf{b}^T\mathbf{b} = 1, \quad (11)$$

where $\times$ denotes the vector product, defined as

$$\mathbf{a}\times\mathbf{b} := (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)^T.$$

Recall that, for $\mathbf{a}, \mathbf{b}, \mathbf{c}$ as in (11), $\mathbf{a} = \mathbf{b}\times\mathbf{c}$ and $\mathbf{b} = \mathbf{c}\times\mathbf{a}$. Using $c_3 = a_1b_2 - a_2b_1$, we obtain $\langle B, S\rangle = a_1 - a_2 + b_1 + b_2 - a_1b_2 + a_2b_1$. Therefore our task is now to prove that the optimum value of the program

$$\begin{array}{ll} \max & f := a_1 - a_2 + b_1 + b_2 - a_1b_2 + a_2b_1 \\ \text{s.t.} & g_1 := a_1^2 + a_2^2 + a_3^2 = 1 \\ & g_2 := b_1^2 + b_2^2 + b_3^2 = 1 \\ & g_3 := a_1b_1 + a_2b_2 + a_3b_3 = 0 \end{array} \quad (12)$$

is at most $2\sqrt{2}-1$; we in fact show that $\max f = 2\sqrt{2}-1$. For this, consider a global maximizer (a, b) to the program (12). Then, the Karush-Kuhn-Tucker conditions have to be satisfied, since the gradient vectors $\{\nabla g_i(a,b) \mid i = 1,2,3\}$ are linearly independent; see, e.g., Theorem 12.1 in [16]. (Here, the gradient vector $\nabla g_i(a,b)$ consists of the partial derivatives with respect to the six variables $a_1, \ldots, b_3$.) That is, there exist scalars $\lambda_1, \lambda_2, \lambda_3$ for which

$$\nabla f(a,b) + \sum_{i=1,2,3} \lambda_i \nabla g_i(a,b) = 0.$$

Equivalently, considering the partial derivatives first with respect to (a_1, a_2, a_3) and then with respect to (b_1, b_2, b_3)

$$\begin{array}{rcl} \begin{pmatrix} 1-b_2 \\ -1+b_1 \\ 0 \end{pmatrix} + 2\lambda_1\mathbf{a} \qquad + \lambda_3\mathbf{b} & = & 0 \\ \begin{pmatrix} 1+a_2 \\ 1-a_1 \\ 0 \end{pmatrix} \qquad + 2\lambda_2\mathbf{b} + \lambda_3\mathbf{a} & = & 0. \end{array}$$

Multiplying the first and the second line by $\mathbf{c}^T = (\mathbf{a}\times\mathbf{b})^T$ (recall that $\mathbf{c} \perp \mathbf{a}, \mathbf{b}$) we get

$$\begin{aligned} 0 &= c_1(1-b_2) + c_2(-1+b_1) = c_1 - c_2 + a_3 \\ 0 &= c_1(1+a_2) + c_2(1-a_1) = c_1 + c_2 + b_3. \end{aligned}$$

Adding (resp. subtracting) these equations yields $2c_1 = -a_3 - b_3$ and $2c_2 = a_3 - b_3$. Squaring these two equations and then adding them gives $2a_3^2 + 2b_3^2 = 4c_1^2 + 4c_2^2$. Since the rows and columns in S are normalized, we get $2(1-c_3^2) = 4(1-c_3^2)$, from which we conclude $c_3^2 = 1$ and, therefore, $a_3 = b_3 = c_1 = c_2 = 0$. This implies $a_1^2 + b_1^2 = 1 = a_1^2 + a_2^2$ and thus $|b_1| = |a_2|$. Similarly one can establish $|a_1| = |b_2|$. On the basis of this observation we distinguish three cases.

1. $a_1 = b_2 = 0$. Then, $|a_2| = |b_1| = 1$ and $f = -a_2 + b_1 + a_2 b_1 \leq 1$.

2. $a_1 \neq 0$ and $a_1 = -b_2$. From $a^T b = 0$ we have $a_1(b_1 - a_2) = 0$, which gives $a_2 = b_1$. Then, $f = a_1 - a_2 + a_2 - a_1 + a_1^2 + a_2^2 = 1$.

3. $a_1 \neq 0$ and $a_1 = b_2$. From $a^T b = 0$ we have $a_1(b_1 + a_2) = 0$, which gives $a_2 = -b_1$. Then, $f = a_1 - a_2 - a_2 + a_1 - a_1^2 - a_2^2 = 2(a_1 - a_2) - 1$, which (under the condition $a_1^2 + a_2^2 = 1$) is clearly maximized by $a_1 = -a_2 = 1/\sqrt{2}$. Therefore, we find $\max f = 2\sqrt{2} - 1$.

Thus, we have shown that the optimum value of the program (12) is equal to $2\sqrt{2} - 1$, which concludes the proof. □

Lemma 3. *Let $f : \{0,1\}^{m_A} \times \{0,1\}^{m_B} \to \{0,1\}$ be a function and K a quantum circuit for f with error probability at most ϵ which uses only gates from* CLIFFORD* *and measurements in the computational basis. Then $C_f^\epsilon \leq 1$.*

Proof. From Lemma 1 we know how two parties, Alice and Bob, can simulate perfect Clifford gates. From Lemma 2 we know how they can jointly simulate the other noisy 1-qubit gates in CLIFFORD*, where they use shared randomness to make sure that they always simulate the same Clifford gate. Thus, Alice and Bob can just simulate the CLIFFORD* circuit for f, using one bit of communication in the end. □

We can now prove an upper bound on the noise in fault-tolerant quantum computation.

Theorem 1. *The set of gates from* CLIFFORD *together with 1-qubit gates with depolarizing noise more than $\hat{\theta}$ and one single-qubit measurement is not sufficient for arbitrary classical computation.*

Proof. The result follows by Lemma 3 and the fact that there are functions with communication complexity greater than 1, for any bounded error. □

In fact we have that none of the functions f with $C_f^\epsilon > 1$ can be computed by CLIFFORD* circuits with error at most ϵ. From Corollary 1 we also get that functions computable by CLIFFORD* are exactly those which can be written as probabilistic mixtures of parity functions.

5 Discussion and extensions

Best gates From the proof of Lemma 2 we see that the rotation matrix S which achieves the optimal value, is

$$\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Multiplying from the right by the Clifford-matrix $diag(1,-1,-1)$ we get a rotation around the z-axis by $\pi/4$. The $\pi/8$-gate

$$T = \begin{pmatrix} \exp(-i\pi/8) & 0 \\ 0 & \exp(i\pi/8) \end{pmatrix}$$

performs a rotation of $\pi/4$ around the z-axis. So, the $\pi/8$-gate and its symmetric versions are the ones which need the most depolarizing noise to be simulated by gates from CLIFFORD.

Worst case noise In Lemma 2 we asked with how much depolarizing noise all 1-qubit unitary gates are equivalent to probabilistic mixtures of Clifford gates. Similarly to [26] one can also ask how much arbitrary noise is needed to make every gate a mixture of Cliffords. More precisely what is the value $\tilde{\theta} = \sup_{U \in SU(2)} p_U$, where p_U is the infimum of all p such that there is a completely positive trace-preserving 1-qubit quantum operation $\mathcal{E}_U$ with the property that the noisy implementation of U

$$U' : \rho \mapsto (1-p)U\rho U^* + p\mathcal{E}_U(\rho)$$

becomes a probabilistic mixture of Clifford operation.

In this section we will provide some bounds on $\tilde{\theta}$. Let $K \in SU(2)$ be any operation that in Bloch representation maps the state X-eigenstate $v_X = (1,0,0)$ to $u = \frac{1}{\sqrt{3}}(1,1,1)$. Note that a probabilistic mixture of 1-qubit Clifford operations $C = \sum_i p_i C_i$ can map v_X only into the octahedron $\mathcal{O}$ spanned by $v_X = (1,0,0)$, $v_Y = (0,1,0)$ and $v_Z = (0,0,1)$ and their negatives $-v_X, -v_Y, -v_Z$ (see also [5]). Note that the state of $\mathcal{O}$ which is closest to u is $\frac{1}{3}(1,1,1) = \frac{1}{\sqrt{3}}u$ and their distance is $||u - 1/\sqrt{3}u||_2 = 1 - \frac{1}{\sqrt{3}}$. The Bloch-state which is furthest away from u is $-u$. All three of these states lie on a line. With this it is clear that the state u_{noise} which needs the smallest noise p, such that $(1-p)u + pu_{noise}$ is inside the octahedron is $-u$ and the optimal p is $\frac{1}{2}(1 - \frac{1}{\sqrt{3}})$. This implies $21\% \approx \frac{1}{2}(1 - \frac{1}{\sqrt{3}}) \leq \tilde{\theta}$.

To get an upper bound, recall that by Lemma 2 for any gate $U \in SU(2)$ the operation

$$U' : \rho \mapsto (1-p)U\rho U^* + p\mathbb{I}/2$$

is a Clifford operation, if $p \geq \hat{\theta}$. Setting $\mathcal{E}_U(\rho) = \frac{1}{3}\left(XU\rho U^* X + YU\rho U^* Y + ZU\rho U^* Z\right)$ and noting that for any 1-qubit density matrix it holds $\mathbb{I}/2 = \frac{1}{4}\left(\rho + X\rho X + Y\rho Y + Z\rho Z\right)$ we can rewrite the action of U' also as

$$U' : \rho \mapsto (1 - \frac{3}{4}p)U\rho U^* + \frac{3}{4}p\mathcal{E}_U(\rho).$$

Thus $\tilde{\theta} \leq \frac{3}{4}\hat{\theta} \approx 34\%$. Note that this is certainly not tight, since all gates, apart from the $\pi/8$-gate (and its symmetric versions), need less than $\hat{\theta}$ depolarizing noise to make it a probabilistic mix of Clifford operations, which implies they need less than $\frac{3}{4}\hat{\theta}$ worst case noise. However, as follows from [26], the worst case noise for the $\pi/8$-gate(s) is only $\frac{1}{2} - \frac{1}{2\sqrt{2}} \approx 15\%$.

We leave it as an interesting open question to determine the precise value of $\hat{\theta}$.

Different noise models The approach we have taken can in principal also be applied to other noise models: For any 1-qubit noise operation $\mathcal{E}$, with Bloch representation $S_{\mathcal{E}}$ we can compute the minimum value θ such that for all rotations $R \in \mathbb{R}^{3\times 3}$ the noisy version $(1-\theta)R + \theta S_{\mathcal{E}}$ is inside the Clifford polytope P (6). However, the actual optimization problems might not be as easy as for depolarizing noise, since depolarizing noise with probability p corresponds to multiplying with $(1-p)$ in Bloch-representation.

In principal, a similar approach might be possible to calculate how well one can approximate arbitrary (unitary) gates given a gate set S other than CLIFFORD* under a certain noise model. If S is not universal, this will give new noise bounds, too.

Classical Co-processing Theorem 1 states that fault tolerant quantum computing is not possible if we have depolarizing noise at least $\hat{\theta} \approx 45\%$ on one qubit gates even if we can use perfect gates from CLIFFORD in our fault tolerant circuit design. Is this optimal? Could it be that with less than $\hat{\theta}$ noise on the single qubit gates and perfect gates from CLIFFORD still no fault tolerant circuit design is possible. We leave this as an open question, but Ben Reichardt [19] pointed out that when we allow perfect classical computation in addition to perfect gates from CLIFFORD and perfect measurements in the computational basis, for any quantum circuit one can build a fault tolerant quantum circuit, that tolerates noise less than $\hat{\theta}$ on single qubit gates. This fault tolerent implementation has only a constant factor slowdown in time.

The argument builds upon magic-state distillation, introduced in [5], and goes as follows. Assume we have at our disposal noisy $\pi/8$-gates T', with depolarizing noise strictly less than $\tilde{\theta}$, i.e. $T'(\rho) = (1-p)T\rho T^* + p\mathbb{I}/2$ with $p < \hat{\theta}$, where T is the perfect $\pi/8$ gate. Then apply T' to the second half of an EPR-pair and measure the observable $Z \otimes Z$, which can be implemented as a measurement in the computation basis with additional gates from CLIFFORD. If the outcome is -1 throw away the state and do the experiment again. If the outcome is $+1$, apply a CNOT from the first to the second qubit, which gives

$$\frac{1}{2}\left(\mathbb{I} + \frac{1-p}{1-p/2}\frac{1}{\sqrt{2}}X + \frac{1-p}{1-p/2}\frac{1}{\sqrt{2}}Y\right) \otimes |0\rangle\langle 0|. \quad (13)$$

Using the result from [20] an arbitrary supply of qubits in the state of the first qubit of (13) can be used to distill magic states in the H-direction, which together with stabilizer operations is sufficient for quantum computation. We do not know if this also holds for other than the $\pi/8$-gate.

Note that this is tight for the $\pi/8$-gate, since stabilizer operations (Cliffords, measurements in the computational basis and classical co-processing)together with $\pi/8$-gates with depolarizing noise $\hat{\theta}$ can be efficiently simulated classically, as follows from our Lemma 2 and the Gottesman-Knill Theorem.

Allowing some perfect unitaries Our threshold theorem says the following. Let f be a function such that it requires more than one bit of communication in order to compute it, when the input bits are partitioned over Alice and Bob. There is no quantum circuit consisting of *perfect* Clifford operations and single qubit gates with noise $\hat{\theta}$ $(\approx 45\%)$ that can compute f. We can strengthen this result to allow a small amount of perfect single qubits as well. Assume that f requires m bits of communication to be computed. There is no quantum circuit that uses perfect Clifford operations, s perfect single qubit gates, and single qubit gates with noise $\hat{\theta}$ that computes f, for $2s+1 < m$. The reason we get this strengthening is because in our simulation, Lemmas 1 and 2, Alice sends to Bob whenever he wants to perform a perfect single qubit gate on some qubit, her classical share a and b of that specific qubit. Bob can now perform the perfect qubit gate on that qubit and they proceed as in Lemma 1 and 2. By the end of the simulation Alice has sent $2s+1$ bits to Bob and he will be able to compute f, contradicting that the communication complexity of f is at least $m > 2s+1$.

Quantum inputs Lemma 1 can actually be extended to the case where Alice and Bob get quantum states as inputs and they are provided with entanglement. It is no problem for Bob to start with a quantum state as an input. For Alice we do the following. We let her teleport her quantum input to Bob bit by bit, using the standard scheme for teleportation (see e.g. [15]). When Alice teleports a qubit, which corresponds to the i-th input qubit of the circuit C to be simulated, she measures two classical bits. Now, if she does not send these to Bob, but rather initializes her a_i, b_i with these bits, Alice and Bob obtain the correct representation for qubits of C as in Lemma 1. Since the inner product function has communication complexity $\Omega(n)$ even in the presence of entanglement [7] we see that Theorem 1 is also true for quantum inputs.

6 Acknowledgements

We thank Nebojša Gvozdenović and Hartwig Bosse and also Troy Lee, Ben Reichardt and Stephanie Wehner for useful discussions and Komei Fukuda for providing the software cdd+ [9]. We also thank Scott Aaronson and Daniel Gottesman for discussions about their work. We are also grateful to the Newton Institute, Cambridge, where this work was started.

References

[1] S. Aaronson and D. Gottesman. Improved simulation of stabilizer circuits. Technical report, arXiv:quant-ph/0406196 v4, 2004.

[2] D. Aharonov and M. Ben-Or. Polynomial simulations of decohered quantum computers. *Proceedings of the 37th IEEE FOCS*, pages 46–55, 1996.

[3] D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error. In *STOC*, pages 176–188, 1997.

[4] G. Brassard, H. Buhrman, N. Linden, A. Methot, A. Tapp, and F. Unger. A limit on nonlocality in any world in which communication complexity is not trivial. *Phys. Rev. Lett.*, 96:250401, 2006.

[5] S. Bravyi and A. Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Phys. Rev. A 71, 022316*, 2005.

[6] A. R. Calderbank and P. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A 52, 2493*, 1996.

[7] R. Cleve, W. van Dam, M. Nielsen, and A. Tapp. Quantum entanglement and the communication complexity of the inner product function. *Selected papers from the First NASA International Conference on Quantum Computing and Quantum Communications*, pages 61–74, 1998.

[8] R. Cleve and J. Watrous. Fast parallel circuits for the quantum fourier transform. In *FOCS*, pages 526–536, 2000.

[9] K. Fukuda. *Software* cdd+. available from http://www.ifor.math.ethz.ch/ fukuda.

[10] D. Gottesman and I. Chuang. Demostrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402:390–393, 1999.

[11] A. Harrow and M. Nielsen. How robust is a quantum gate in the presence of noise? *Phys. Rev. A 68, 012308*, 2003.

[12] E. Knill. Quantum computing with realistically noisy devices. *Nature 434, 39-44*, 2005.

[13] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[14] G. Nebe, E. M. Rains, and N. J. A. Sloane. The invariants of the clifford groups. *Designs, Codes and Cryptography*, 24(99), 2001. math.CO/0001038.

[15] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[16] J. Nocedal and S. Wright. *Numerical Optimization*. Springer Press, New York, 1999.

[17] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Physical Review Letters*, 86:5188, 2001.

[18] A. Razborov. An upper bound on the threshold quantum decoherence rate. *Quantum Information and Computation*, 4(3):222–228, 2004.

[19] B. Reichardt. personal communicaton. 2005.

[20] B. Reichardt. Quantum universality from magic states distillation applied to css codes. In *Quantum Information Processing*, volume 4, pages 251 – 264. Springer, Netherlands, 2005.

[21] P. Shor. Scheme for reducing decoherence in quantum memory. *Phys. Rev. A 52, 2493*, 1995.

[22] P. Shor. Fault-tolerant quantum computation. In *Proc. 37th Annual Symposium on Foundations of Computer Science*, pages 56–65. IEEE Computer Society Press, 1996.

[23] R. Solovay. talk at mathematical sciences research institute.

[24] A. Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett. 77, 793*, 1996.

[25] M. Szegedy. Functions with bounded symmetric communication complexity, programs over commutative monoids, and acc. *Journal of Computer and System Sciences*, 47:405–423, 1993.

[26] S. Virmani, S. Huelga, and M. Plenio. Classical simulatability, entanglement breaking, and quantum computation thresholds. Technical report, quant-ph/0408076, 2004.