One-Way Quantum Communication Complexity

Dominic Moylett*

Quantum Engineering Centre for Doctoral Training,
University of Bristol

March 24, 2016

1 Introduction

Ever since the idea of quantum computing was first suggested, one of the most important questions has been "What advantages would a quantum computer have over classical computers?" In the decades that followed, quantum computers have shown many benefits, from factoring integers [Sho97] to searching data [Gro96] and even secure communication [BB14]. Given that quantum computers have shown such advantages for a single computer setting, a natural question is can they show advantages in distributed computation?

We can evaluate the advantage of distributed quantum computation via communication complexity. Communication complexity is a very powerful model – enough to earn its creator the first Knuth Prize [1] – and has proven lower bounds in a wide variety of models. It has also led to applications in different forms of computation, from secure multi-party computation to data streaming. But finding separations is not easy, and Holevo's theorem has from the beginning shown us that there are problems which cannot be solved more efficiently via quantum communication.

We will explore communication complexity in this report, focusing specifically on quantum advantages in the one-way communication setting. Section 2 offers a brief overview of communication complexity, and an explanation of Holevo's theorem. Section 3 provides examples of problems for which a separation between quantum and classical communication is either proven or conjectured. We conclude in Section 4 with a summary of the results and some open questions.

2 Communication Complexity

Communication complexity was developed by Yao in 1979 as an analysis of distributed computing [Yao79]. Under this model, our computation is done between two parties, commonly referred to as Alice and Bob, with inputs $x, y \in \{0,1\}^n$, respectively. The two parties do not know each other's inputs; Alice does not know y and Bob does not know x. Alice and Bob then exchange a series of messages between each other through a known protocol, with their aim being to jointly compute $f(x, y)$. The question communication complexity asks is how many bits of information do Alice and Bob need to communicate in order for them. We offer a sketch of a typical communication complexity problem in Figure 1.

A naïve solution to any communication complexity problem uses n bits of communication: Alice sends x to Bob and Bob computes $f(x, y)$, possibly sending the result back to Alice. So the question is "Can we compute $f(x, y)$ with $o(n)$ bits of communication?"

One of the key benefits to communication complexity is that it is a useful model for proving lower bounds, and there are a number of established techniques for doing so. Lower bounds in communication complexity have also led to bounds in the data streaming model of computation[2] such as in [GKK+07]

*dominic.moylett@bristol.ac.uk

[1] http://www.sigact.org/Prizes/Knuth/1996.html.

[2] See [Mut05] for an explanation of the data streaming model.

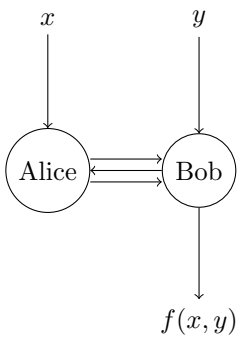

Figure 1: An example of communication complexity. Alice and Bob begin computation with bit strings x and y. They then exchange a series of messages based on a pre-determined protocol. Computation completes when Bob outputs $f(x, y)$.

& [VY11], as well as lower bounds for quantum and classical circuit depth [Ker09, Yao93]. Beyond the computation of lower bounds, other applications to communication complexity include classical security [GKK+07] and secure multiparty computation in both classical and quantum settings [FY92, DPP14, KLGR14].

In one-way communication complexity, Alice and Bob again receive respective inputs $x, y \in \{0, 1\}^n$ with the aim being for Bob to output $f(x, y)$ for some function f. The difference now is that only one message is allowed to be sent from Alice to Bob, and Bob cannot send any messages to Alice.

2.1 Quantum Communication Complexity

In quantum communication complexity, first proposed in 1993 by Yao [Yao93], we also have parties Alice and Bob given inputs x and y, respectively, who have to compute $f(x, y)$ for some function f. But now our two parties can send quantum messages between each other, and the complexity is measured in the number of qubits that are communicated.

Another model by Cleve and Buhrman [CB97] does not use the physical sending of data between the parties, but instead uses the measurement of entangled particles for remote computation. While our focus for this report however will be on the former, it is possible to simulate communicating l individual qubits by use of l entangled pairs and $2l$ classical bits of communication, via quantum teleportation [BBC+93]. Likewise, it is possible to simulate this model by the ability to communicate quantum states: Alice generates entangled pairs of qubits and sends one of the pair to Bob. This and other models are briefly discussed in Appendix E.

Because of the well-established techniques for lower-bounding classical communication complexity, a common area of research is to see if there are any problems which quantum protocols can solve by communicating fewer qubits than the number of classical bits required by lower bounds. Finding quantum protocols that beat classical lower bounds can give us intuition on the difference between quantum and classical encoding of data. These protocols will be the focus of Section 2.

2.2 Holevo's Theorem

Even before Yao's work on the formalisation of Communication Complexity, it was clear that quantum communication would not be able to speed up some problems. A result by Holevo [Hol73] gave an upper limit on the amount of information that can be extracted from a quantum state. A consequence of Holevo's theorem was that while a message of n bits could be compressed into a message of less than n qubits, the same message could not be retrieved from said qubits with absolute certainty. We phrase this as a communication complexity problem below.

Problem COMMUNICATION

Alice's input: $x \in \{0, 1\}^n$.
Bob's input: none.
Bob's output: x.

Theorem 2.1. *Any quantum protocol for COMMUNICATION that uses $O(2^m)$ qubits of communication succeeds with probability at most 2^{m-n}.*

We can prove Theorem 2.1 by Holevo's theorem, but a later proof by Nayak [Nay99] explains this result explicitly for encoding classical data in quantum data. This proof is given in Appendix A.

The only exception to this is if Alice and Bob have shared entangled qubits between each other prior to computation, in which case superdense coding can be used to communicate two bits by physically sending one qubit [BW92]. On the other hand, if Alice and Bob share prior entanglement and can communicate classically, they can communicate single qubits by physically sending pairs of bits [BBC+93], so we can only save at most a factor of 1/2 [Bra01].

Holevo's Theorem has in turn led to communication complexity problems which cannot be solved more efficiently with quantum communication. One example of this is the Inner Product problem:

Problem INNER-PRODUCT
ALICE'S INPUT: $x \in \{0,1\}^n$.
BOB'S INPUT: $y \in \{0,1\}^n$.
BOB'S OUTPUT: $\sum_{i=0}^{n-1} x_i.y_i(\text{mod}2)$.

Cleve et al. [CvDNT13] showed that if we were able to solve this problem in less than $O(n)$ qubits of communication, then it would be possible for Alice to communicate x to Bob in $o(n)$ qubits, thus violating Holevo's theorem.

3 Separations Between Quantum and Classical Communication Complexity

While it is clear from Holevo's theorem that we cannot solve all problems in fewer qubits than we can in classical bits, there are a number of problems which can be solved by communicating fewer qubits than classical bits. Some of the significant results are described in this section.

3.1 Distributed Deutsch-Jozsa

The Distributed Deutsch-Jozsa problem, first introduced by Buhrman, Cleve and Wigderson [BCW98], was the first example of an exponential separation between quantum and classical computation complexity. Prior to this, the largest gaps between classical and quantum communication were constant factors [CB97]. The problem, and its quantum protocol as described in Buhrman et al. [BCMdW10], are stated as follows, where $d(x,y)$ is the Hamming distance between x and y and $U_x = \sum_{i=0}^{n-1}(-1)^{x_i}|i\rangle\langle i|$ for $x \in 0,1^n$:

Protocol DISTRIBUTED DEUTSCH-JOZSA
ALICE'S INPUT: $x \in \{0,1\}^n$.
BOB'S INPUT: $y \in \{0,1\}^n$.
PROMISE: $x = y$ or $d(x,y) = n/2$.
BOB'S OUTPUT: 0 iff $x = y$, 1 otherwise.
1 Alice starts computation in the $|0\rangle^{\otimes \log n}$ state and applies $U_x H^{\otimes \log n}$ to get the state $|\psi\rangle$.
2 Alice sends $|\psi\rangle$ to Bob.
3 Bob applies $H^{\otimes \log n} U_y$ and measures in the computational basis.
4 Bob outputs 0 if they measure $|0\rangle^{\otimes \log n}$ and 1 otherwise.

The state Alice sends to Bob is $|\psi\rangle = \frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}(-1)^{x_i}|i\rangle$. When Bob applies U_y to this state, they end up with the state $\frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}(-1)^{x_i}(-1)^{y_i}|i\rangle$. We'll now consider this state for two situations: one where $x = y$ and one where they differ in $n/2$ places, denoted $|\psi_0\rangle$ and $|\psi_1\rangle$, respectively.

Note that if $x = y$ then $(-1)^{x_i}(-1)^{y_i} = 1$, so $|\psi_0\rangle = \frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}|i\rangle$. $|\psi_1\rangle$ on the other hand will have exactly $n/2$ amplitudes which are positive, and $n/2$ amplitudes which are negative. From this

we can see that $\langle\psi_0|\psi_1\rangle = \frac{n/2-n/2}{n} = 0$, so the states can be distinguished from one another upon measurement. Finally, Bob's application of $H^{\otimes \log n}$ converts their state to the computational basis. Note that $H|\psi_o\rangle = |0\rangle^{\otimes \log n}$, so measuring this state will indicate that $x = y$ and any other state will indicate that $d(x, y) = n/2$.

So we have been able to deterministically solve this problem with $O(\log n)$ qubits of communication between Alice and Bob. How much classical communication is required to solve it? Buhrman et al.[BCMdW10] used the work of Buhrman, Cleve and Wigderson [BCW98] to give the following lower bound:

Theorem 3.1. *The Distributed Deutsch-Jozsa Function requires $\Omega(n)$ bits of classical communication to be solved deterministically.*

A summary of this proof is in Appendix B. Thus we have an exponential gap between deterministic classical communication and quantum communication for the Distributed Deutsch-Jozsa problem. However, this quantum advantage is lost when compared to randomised classical communication: Alice picks a random index $i \in \{0, ..., n-1\}$ and sends (i, x_i) to Bob, who outputs 0 if $x_i = y_i$ and 1 otherwise. This protocol has error 1/2, which can be reduced by repetition. Overall communication is $O(\log n)$ bits.

3.2 Subgroup Membership

Another disadvantage with the Distributed Deutsch-Jozsa problem is that it is a partial function, in that it requires a promise. This begs the question: Are there any functions which do not rely on promises – better known as total functions – but still have a quantum advantage?

This is an open question, but some progress has been made. Aaronson et al. [AGRT09] explored the problem of subgroup membership as potentially yielding a quantum advantage for a total function. Their protocol for solving subgroup membership, based on an earlier algorithm by Watrous [Wat00], is described below, where G is a group, $\mathscr{H}_G$ is the set of all subgroups of G, $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $U_{G,y} = \sum_{g\in G} |yg\rangle\langle g|$ for some $y \in G$:

Protocol SUBGROUP-MEMBERSHIP

ALICE'S INPUT: $H \in \mathscr{H}_G$.
BOB'S INPUT: $y \in G$.
BOB'S OUTPUT: 1 if $y \in H$, 0 otherwise.

1 Alice prepares the state $|H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h\in H} |h\rangle$, which is then sent to Bob.
2 Bob prepends an ancilla qubit to get $|H\rangle|+\rangle$ and applies $U_{G,y}$ controlled on the final qubit.
3 Bob applies a Hadamard to the second qubit and measures in the computational basis.

Any element $h \in H$ can be described in $\lceil \log |G| \rceil$ bits, so Alice needs to send $O(\log |G|)$ qubits to Bob in order to perform this protocol.

After line 2, Bob has the state $\frac{1}{\sqrt{2}}(|H\rangle|0\rangle + |yH\rangle|1\rangle)$, where $|yH\rangle = \frac{1}{\sqrt{|H|}} \sum_{h\in H} |yh\rangle$. If $y \in H$ then $U_{G,y}|H\rangle|+\rangle = |H\rangle|+\rangle$, so applying a Hadamard to the final qubit makes the final qubit $|0\rangle$. A measurement in the computational basis will thus always yield 0.

On the other hand, if $y \notin H$ then $\langle yH|H\rangle = 0$, so applying the Hadamard produces

$$\frac{1}{2}((|H\rangle + |yH\rangle)|0\rangle + (|H\rangle - |yH\rangle)|1\rangle).$$

Measuring the ancilla qubit will result in $|0\rangle$ – and thus a false positive – with probability 1/2. We can repeat this protocol a constant number of times to reduce this error.

Classically, it has been shown by Miltersen et al. [MNSW98] that any one-way randomised protocol with completeness probability of 1/2 and perfect soundness must use $\Omega(\log^2 |G|)$ bits of communication. Aaronson et al. have conjectured that this bound still holds even in a bounded-error setting.

So if we can find a total subgroup membership function that cannot classically be solved using $O(\log |G|)$ bits of communication, then we will have a quadratic gap for a total function. The problem is that there are many subgroups which can be solved classically. For example, Aaronson et al. gave the following theorem for normal subgroups:

Theorem 3.2. *For normal subgroup $H \in \mathscr{H}_G, y \in G$ there is an $O(\log |G|)$ protocol which outputs with probability 1 if $y \in H$ and 0 with probability at least 1/2 if $y \notin H$.*

To explain this classical protocol, we first need to describe complex characters. For any complex field representation ρ of G, the complex character χ is the function $\chi(g) = \mathrm{tr}(\rho(g))\ \forall\ g \in G$. For any two elements $g, g' \in G$, we have the relation that $\chi(gg'g^{-1}) = \chi(g)$. $\chi(1)$ is the dimension of ρ, and the kernel of χ, denoted $\ker(\chi) = \{g \in G | \chi(g) = chi(1)\}$. If ρ is irreducible then χ is an irreducible character, and the set of irreducible complex characters of G is denoted $\mathrm{Char}(g)$. If H is a normal subgroup of G, then let

$$\Lambda_H = \{\chi \in \mathrm{G} | H \subseteq \ker(\chi)\}.$$

Given this definition we note that

$$\sum_{\chi \in \Lambda_H} [\chi(1)]^2 = \frac{|G|}{|H|}.$$

We now describe the protocol from Aaronson et al. [AGRT09]:

Protocol NORMAL-SUBGROUP-MEMBERSHIP

ALICE'S INPUT: $H \in \mathscr{H}_G$ such that H is normal.
BOB'S INPUT: $y \in G$.
BOB'S OUTPUT: 1 if $y \in H$, 0 otherwise.

1 Alice chooses a random complex character μ such that $H \subseteq \ker(\mu)$ with probability $\frac{|H|}{|G|}[\mu(1)]^2$.
2 Alice sends μ to Bob.
3 Bob outputs 1 if $y \in \ker(\mu)$ and 0 otherwise.

Proof of Theorem 3.2. If $y \in H \subseteq \ker(\mu)$ then $\mu(y) = \mu(1)$ and thus the protocol will succeed. So the protocol fails if Alice picks a character μ such that $y \in \ker(\mu) \setminus H$. We shall call the set of all characters that would cause an error B. Based on Alice's selection probability, the probability of an error is thus

$$\frac{|H|}{|G|} \sum_{\chi \in B} [\chi(1)]^2.$$

To analyse this further, we will look at the smallest normal subgroup of G which contains both H and y. We will call this subgroup K, which can be stated formally as $K = \{gkg^{-1} | g \in G, k \in H \cup \{y\}\}$. Note that $|K| \geq 2|H|$, as $h \in K$ and $hyh^{-1} \in K$, but $hyh^{-1} \notin H\ \forall\ h \in H$.

The final point we will need is that $B \subseteq \Lambda_K$. Consider a character $\chi \in B$. Because χ could be selected by Alice, we know that $H \subseteq \ker(\chi)$. We also know that $y \in \ker(\chi)$ because $\chi \in B$. We see from this that any other element in K must also be in $\ker(\chi)$, as $\chi(gkg^{-1}) = \chi(k)\ \forall\ g \in G, k \in H \cup \{y\}$. Thus $K \subseteq \ker(\chi)$ and we can conclude that $\chi \in \Lambda_K$.

Using the above points, we can show that

$$\frac{|H|}{|G|} \sum_{\chi \in B} [\chi(1)]^2 \leq \frac{|H|}{|G|} \sum_{\chi \in \Lambda_K} [\chi(1)]^2 = \frac{|H|}{|G|} \frac{|G|}{|K|} \leq \frac{|H|}{2|H|} = \frac{1}{2}.$$

Thus this protocol fails with probability at most 1/2. Finally, we note that is is possible to describe μ in $\lceil \log |G| \rceil$ bits, as $\mu \in \Lambda_H$ and $|\lambda_H| \leq |Char(G)| \leq |G|$. □

Other groups also proven to be solvable in $(O \log |G|)$ bits of communication include groups with small modular or representations and groups with small complex representations [AGRT09]. So already we have three cases of groups which don't have a separation.

3.3 Hidden Matching

A matching M is a sequence of k pairs $(i_1, j_1), (i_2, j_2), ..., (i_k, j_k) \in [n]^2$. If no integer appears in any pair more than once, then the pairs are disjoint. A matching is said to be partial the pairs are disjoint

but some indexes do not appear in any pair, and perfect if each integer appears in exactly one pair. A matching M together with an n-bit string x define a k-bit string as follows:

$$Mx = (x_{i_1} \oplus x_{j_1}), (x_{i_2} \oplus x_{j_2}), ...(x_{i_k} \oplus x_{j_k})$$

The Hidden Matching problem was first investigated in terms of one-way communication complexity in [BYJK08]. The problem and its quantum protocol are defined below, where $\mathcal{M}_n$ is the set of perfect matchings on n-node graphs:

Protocol HIDDEN-MATCHING

ALICE'S INPUT: $x \in \{0,1\}^n$.
BOB'S INPUT: $M \in \mathcal{M}_n$.
BOB'S OUTPUT: $(i, j, x_i \oplus x_j)$ for some $(i, j) \in M$.

1 Alice prepares the state $|x\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i} |i\rangle$, which is then sent to Bob.
2 Bob measures $|x\rangle$ in the $\{\frac{1}{\sqrt{2}}(|i\rangle \pm |j\rangle) | (i, j) \in M\}$ basis and gets the state $|\psi_{i\pm j}\rangle$.
3 Bob returns (i, j, b), where $b = 0$ if $|\psi_{i\pm j}\rangle = \frac{\pm 1}{\sqrt{2}}(|i\rangle + |j\rangle)$ and $b = 1$ if $|\psi_{i\pm j}\rangle = \frac{\pm 1}{\sqrt{2}}(|i\rangle - |j\rangle)$.

First, note that since M is a perfect matching, each index $i \in \{0, ..., n-1\}$ occurs in exactly one (i, j)-pair in M. Thus the basis of step 2 is orthonormal and can be measured. The probability of measuring $|\psi_{i\pm j}\rangle$ is

$$P(i \pm j) = \left| \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} (-1)^{x_k} \frac{1}{\sqrt{2}} (|i\rangle \pm |j\rangle) \right|^2 = \frac{|(-1)^{x_i} \pm (-1)^{x_j}|^2}{2n}$$

.

If $x_i \oplus x_j = 0$ then $x_i = x_j$, so $P(i - j) = \frac{1}{2n}(1 - 1) = 0$ and thus Bob will measure the state $|\psi_{i+j}\rangle$. Likewise, if $x_i \oplus x_j = 1$ then Bob will measure the state $|\psi_{i-j}\rangle$. Thus this quantum protocol will always find a pair, though a specific pair has probability $\frac{1}{2n}$.

As for classical one-way protocols, the following theorem was proven in [BYJK08]:

Theorem 3.3. *A random one-way protocol for the Hidden Matching problem with deterministic Alice requires $O(\sqrt{n})$ bits of communication.*

This proof is summarised in Appendix C.

Unfortunately, this is not an example of a total function that separates quantum and classical one-way communication complexity. This is because the hidden matching problem has a one-to-many mapping: Bob is allowed to output any $(i, j, x_i \oplus x_j)$ tuple under the condition that $(i, j) \in M$, thus there are $n/2$ acceptable outputs. This makes the hidden matching problem not a function, but a relation. A separation for relations does not necessarily prove a separation for functions.

3.4 α-Matching

α-Matching is a promise function related to the hidden matching problem. In α-matching, the matching given to Bob consists of αn disjoint pairs. If $\alpha = 1/2$ then the matching is perfect like before, and if $\alpha < 1/2$ then the matching is partial. The problem and its quantum protocol from [GKK+07] are described as follows, where $\mathcal{M}_{\alpha n}$ is the set of all α-matchings, and $\bar{x}$ is the bitwise NOT of the string x:

Protocol α-MATCHING

ALICE'S INPUT: $x \in \{0,1\}^n$.
BOB'S INPUT: $M \in \mathcal{M}_{\alpha n}, w \in \{0,1\}^{\alpha n}$.
PROMISE: $w = Mx$ or $\bar{w} = Mx$.
BOB'S OUTPUT: 0 if $w = Mx$, 1 otherwise.

1 Bob adds (i, j) pairs to their matching to get perfect matching M'.
2 Alice and Bob run the HIDDEN-MATCHING protocol on (x, M') to get tuple (i, j, b).
3 If $(i, j) \notin M$ then Bob aborts.
4 Bob outputs 0 if $w_k = b$ and 1 otherwise.

First note that Bob only needs one bit of Mx to figure out which of the two cases holds because of the promise. So if we can find one pair (x_i, x_j) such that $(i, j) \in M$, then we have solved this problem. Since the pairs in M are disjoint, we can pad M with extra pairs to get a reduction to the Hidden Matching problem.

The only case where we will fail is if the hidden matching protocol returns $(i, j, x_i \oplus x_j)$ such that $(i, j) \notin M$. We can reduce this probability by repeating the protocol $O(k/\alpha)$ times for some $k \ll \alpha n$ to get a high probability of successfully finding one bit of Mx with $O(\log n)$ bits of communication.

So how well can we solve this problem with classical communication? Gavinsky et al. [GKK+07] proved the following:

Theorem 3.4. *A random one-way protocol for α-Matching problem with deterministic Alice requires $O(\sqrt{n/\alpha})$ bits of communication.*

Proof. We can prove this by asking how much Bob can learn if Alice sends them c bits, where $1 \leq c \leq \gamma\epsilon\sqrt{n/\alpha}$, for some $\epsilon > 0, \alpha \in (0, 1/4]$ and a carefully chosen value of γ which will be used later.

If Alice's message is deterministic based on x, then the c bits Alice sends allows Bob to define a set $A \subseteq \{0,1\}^n$ such that $|A| \geq 2^{n-c}$ of bit strings from which Bob must be able to guess Alice's input x. The probability of Bob succeeding then is the probability of him picking a bit string $x' \in A$ such that $Mx = Mx'$. We define the probability of Bob finding a string x such that $Mx = z$ for some k-bit string z as

$$p_M(z) = \frac{|\{x \in A | Mx = z\}|}{|A|}.$$

Gavinsky et al. showed that as Alice's message becomes shorter than $O(\sqrt{n/\alpha})$, this probability distribution becomes arbitrarily close to the uniform distribution. They start by looking at the expected value of the squared total variational distance between the probability distribution and the uniform distribution:

$$\mathbb{E}_M[\|p_M - U\|_{tvd}^2] = \mathbb{E}_M[(\sum_{z \in \{0,1\}^{\alpha n}} |p_M(z) - U(z)|)^2]$$

This can then be re-written as a squared inner product:

$$2^{2\alpha n}\mathbb{E}_M\left[\left(\frac{1}{2^{\alpha n}}\sum_{z \in \{0,1\}^{\alpha n}} 1 * |p_M(z) - U(z)|\right)^2\right] = 2^{2\alpha n}\mathbb{E}_M[\langle 1, |p_M(z) - U(z)|\rangle^2]$$

where 1 is the function $1(z) = 1 \ \forall \ z \in \{0,1\}^{\alpha n}$ and $\langle f, g\rangle = \frac{1}{2^{\alpha n}}\sum_{z \in \{0,1\}^{\alpha n}} f(z)g(z)$ is the inner product. By the Cauchy-Schwarz inequality, we can bound this value:

$$2^{2\alpha n}\mathbb{E}_M[\langle 1, |p_M(z) - U(z)|\rangle^2] \leq 2^{2\alpha n}\mathbb{E}_{\mathbb{M}}[\|1\|_2^2\|p_M - U\|_2^2] = 2^{2\alpha n}\mathbb{E}_M[\|p_M - U\|_2^2]$$

where $\|f\|_2^2 = \langle f, f\rangle$ is the l_2-norm. We then use Parseval's identity and note that $\widehat{U}(z) = \widehat{f}(z)$ if $z = 0$ and $\widehat{U}(z) = 0$ otherwise.

$$2^{2\alpha n}\mathbb{E}_{\mathbb{M}}[\|p_M - U\|_2^2] = 2^{2\alpha n}\mathbb{E}_M\left[\sum_{z \in \{0,1\}^{\alpha n}} (\widehat{p_M}(z) - \widehat{U}(z))^2\right] = 2^{2\alpha n}\mathbb{E}_M\left[\sum_{z \in \{0,1\}^{\alpha n} \setminus \{0^{\alpha n}\}} \widehat{p_M}(z)^2\right]$$

where $\widehat{f}(z) = \sum_{y \in \{0,1\}^{\alpha n}} f(y)(-1)^{x.y}$ is the Fourier transform of f and . is the dot product. Now let us look at the fourier transform of p_M:

$$
\begin{aligned}
\widehat{p_M}(z) &= \frac{1}{2^{\alpha n}} \sum_{s \in \{0,1\}^{\alpha n}} p_M(s)(-1)^{s.z} \\
&= \frac{1}{|A| 2^{\alpha n}} \sum_{s \in \{0,1\}^{\alpha n}} |\{y \in A | My = s\}|(-1)^{s.z} \\
&= \frac{1}{|A| 2^{\alpha n}} \sum_{s \in \{0,1\}^{\alpha n}} |\{y \in A | My = s\}|(-1)^{My.z}
\end{aligned}
$$

Note that for any given $y \in A$, there is exactly one $s \in \{0,1\}^{\alpha n}$ such that $My = s$. Because we are summing over all αn-bit strings, we are essentially summing over all $y \in A$:

$$
\begin{aligned}
\widehat{p_M}(z) &= \frac{1}{|A| 2^{\alpha n}} \sum_{s \in \{0,1\}^{\alpha n}} |\{y \in A | My = s\}|(-1)^{My.z} \\
&= \frac{1}{|A| 2^{\alpha n}} \sum_{y \in A} (-1)^{My.z} \\
&= \frac{1}{|A| 2^{\alpha n}} \sum_{y \in A} (-1)^{y.M^T z} \\
&= \frac{1}{|A| 2^{\alpha n}} \sum_{y \in \{0,1\}^n} 1_A(y)(-1)^{y.M^T z} \\
&= \frac{2^n}{|A| 2^{\alpha n}} \widehat{1_A}(M^T z)
\end{aligned}
$$

where 1_A is the indicator function of A. Plugging this back into our formula, we find that

$$
2^{2\alpha n} \mathbb{E}_M \left[\sum_{z \in \{0,1\}^{\alpha n} \setminus \{0^{\alpha n}\}} \widehat{p_M}(z)^2 \right] = \frac{2^{2n}}{|A|^2} \mathbb{E}_M \left[\sum_{z \in \{0,1\}^{\alpha n} \setminus \{0^n\}} \widehat{1_A}(M^T z)^2 \right].
$$

Next, we move from summing over $\alpha - n$ bit strings to summing over n-bit strings, and note that the set $V = |\{s \in \{0,1\}^{\alpha n} | M^T s = v\}| \leq 1 \ \forall \ v \in \{0,1\}^n$. This means that our expectation over M is a probability:

$$
\begin{aligned}
\frac{2^{2n}}{|A|^2} \mathbb{E}_M \left[\sum_{z \in \{0,1\}^{\alpha n} \setminus \{0^n\}} \widehat{1_A}(M^T z)^2 \right] &= \frac{2^{2n}}{|A|^2} \mathbb{E}_M \left[\sum_{v \in \{0,1\}^n \setminus \{0^n\}} |\{s \in \{0,1\}^{\alpha n} | M^T s = v\}| \widehat{1_A}(v)^2 \right] \\
&= \frac{2^{2n}}{|A|^2} \sum_{v \in \{0,1\}^n \setminus \{0^n\}} \Pr_M[\exists \ s \in \{0,1\}^{\alpha n} \text{ s.t. } M^T s = v] \widehat{1_A}(v)^2
\end{aligned}
$$

Now we need to consider what this probability over M is. Suppose the Hamming weight of v is $|v| = k$ for even k[3]. Lemma 5 of [GKK+07] showed that the fraction of α-matchings which satisfy the probability is equal to the number of αn-bit strings with Hamming weight $k/2$ over the number of n-bit strings with Hamming weight k:

$$
\frac{2^{2n}}{|A|^2} \sum_{v \in \{0,1\}^n \setminus \{0^n\}} \Pr_M[\exists \ s \in \{0,1\}^{\alpha n} \text{ s.t. } M^T s = v] \widehat{1_A}(v)^2 = \frac{2^{2n}}{|A|^2} \sum_{\text{even } k=2}^{2\alpha n} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \sum_{v : |v| = k} \widehat{1_A}(v)^2
$$

At this point, we are going assume that $k \leq 4c$ and let $\delta = k/4c$. Because $\delta \in (0, 1]$, we can apply the Kahn, Kalai and Linial inequality [KKL88]:

[3] The Hamming weight of $M^T s$ is even and at most $2\alpha n$.

$$\sum_{\text{even } k=2}^{4c-2} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \frac{2^{2n}}{|A|^2} \sum_{v:|v|=k} \widehat{1_A}(v)^2 \leq \sum_{\text{even } k=2}^{4c-2} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \frac{2^{2n}}{|A|^2} \frac{1}{\delta^k} \left(\frac{|A|}{2^n}\right)^{2/(1+\delta)}$$
$$= \sum_{\text{even } k=2}^{4c-2} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \frac{1}{\delta^k} \left(\frac{2^n}{|A|}\right)^{2\delta/(1+\delta)}$$

We bound this value by the following inequalities: $|A| \geq 2^{n-c}$, $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$ and $c \leq \gamma\epsilon\sqrt{n/\alpha}$:

$$\begin{aligned}
\sum_{\text{even } k=2}^{4c-2} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \frac{1}{\delta^k} \left(\frac{2^n}{|A|}\right)^{2\delta/(1+\delta)} &\leq \sum_{\text{even } k=2}^{4c-2} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \frac{1}{\delta^k} \left(\frac{2^n}{|A|}\right)^{k/2c} \\
&\leq \sum_{\text{even } k=2}^{4c-2} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \left(\frac{4\sqrt{2}c}{k}\right)^k \\
&\leq \sum_{\text{even } k=2}^{4c-2} \frac{(2e\alpha n)^{k/2}}{(n/k)^k} \left(\frac{4\sqrt{2}c}{k}\right)^k \\
&\leq \sum_{\text{even } k=2}^{4c-2} \left(\frac{64e\alpha c^2}{kn}\right)^{k/2} \leq \sum_{\text{even } k=2}^{4c-2} \left(\frac{64e\gamma^2\epsilon^2}{k}\right)^{k/2}
\end{aligned}$$

A small value for γ will ensure that this sum is less than $\epsilon^2/2$.

We now need to bound the sum for values of $k \geq 4c$. For this, we note that the term $\binom{\alpha n}{k/2}/\binom{n}{k}$ decreases as k increases:

$$\begin{aligned}
\frac{\binom{\alpha n}{k/2-1}/\binom{n}{k-2}}{\binom{\alpha n}{k/2}/\binom{n}{k}} &= \frac{(k/2)!(\alpha n - k/2)!}{(k/2-1)!(\alpha n - k/2+1)!} \frac{(k-2)!(n-k+2)!}{k!(n-k)!} \\
&= \frac{k/2}{\alpha n - k/2 + 1} \frac{(n-k+1)(n-k+2)}{k(k-1)} \\
&= \frac{(n-k+2)(n-k+1)}{(2\alpha n - k + 2)(k-1)} \geq \frac{n-k+1}{k-1} \geq 1
\end{aligned}$$

Plugging this in and recalling that $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$, we find that

$$\begin{aligned}
\frac{2^{2n}}{|A|^2} \sum_{\text{even } k=4c}^{2\alpha n} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \sum_{v:|v|=k} \widehat{1_A}(v)^2 &\leq \frac{2^{2n}}{|A|^2} \sum_{\text{even } k=4c}^{2\alpha n} \frac{\binom{\alpha n}{2c}}{\binom{n}{4c}} \sum_{v:|v|=k} \widehat{1_A}(v)^2 \\
&= \frac{2^{2n}}{|A|^2} \left(\frac{8e\alpha c}{n}\right)^{2c} \sum_{\text{even } k=4c}^{2\alpha n} \sum_{v:|v|=k} \widehat{1_A}(v)^2.
\end{aligned}$$

Now we note by Parseval's identity that $\sum_{v\in\{0,1\}^n} \widehat{1_A}(v)^2 = \|1_A\|_2^2 = |A|2^{-n}$, and recall that $|A| \geq 2^{n-c}$ and $c \leq \gamma\epsilon\sqrt{n/\alpha}$:

$$\frac{2^{2n}}{|A|^2} \left(\frac{8e\alpha c}{n}\right)^{2c} \sum_{\text{even } k=4c}^{2\alpha n} \sum_{v:|v|=k} \widehat{1_A}(v)^2 \leq \frac{2^n}{|A|} \left(\frac{8e\alpha c}{n}\right)^{2c} \leq 2^c \left(\frac{8e\alpha c}{n}\right)^{2c} \leq \left(\frac{8\sqrt{2}e\gamma\epsilon\sqrt{n}}{\sqrt{\alpha}}\right)^{2c}$$

Again, this equation can be bounded by $\epsilon^2/2$ for small enough γ, as $\alpha/n < 1$, $c \geq 1$. By adding these two cases together, we find that the squared total variation distance is at most ϵ^2. To conclude, we use Jensen's inequality:

$$(\mathbb{E}_M[\|p_M - U\|_{tvd}])^2 \leq \mathbb{E}_M[\|p_M - U\|_{tvd}^2] \leq \epsilon^2$$

We conclude that if Alice sends fewer than $\sqrt{n/\alpha}$ bits of classical data to Bob then this distribution is indistinguishable form uniform. □

One potential weakness with this proof is the use of *gamma*, which is to essentially force the message to be small enough that this proof holds. It would be interesting to see to what extent we can bound this relation if $\gamma = 1$.

A more general version of the problem was explored by Verbin and Yu in 2011 [VY11], where showed that t-dimensional hypermatching requires $\Omega(n^{1-1/t})$ bits of classical communication. Note that $t = 2$ is consistent with Bar-Yossef, Jayram and Kerenidis' results for perfect matching [BYJK08].

Another partial function related to both Subgroup Membership and α-matching is Permutation Invariance [Mon11], which is discussed in Appendix D.

4 Conclusion

Communication complexity has proven itself to be a powerful tool in theoretical computer science, from its origins in distributed computation to lower bounds in the circuit depth & streaming models alongside applications in cryptography. It has also proven over the last two and a half decades to be useful for finding separations between classical and quantum computing.

We have presented in this report some of the one-way communication relations and partial functions for which the communication of bits is significantly worse than qubits. It is important to note that a number of other grand separations have also been proven which we have not been able to cover in this essay: Gavinsky [Gav08] showed a relation which had an exponential separation between one-way quantum communication and two-way randomised classical communication. Regev and Klartag [RK11] later showed a partial function with the same separation.

While we have focused on one-way communication complexity as well as some of two-way in this conclusion, there are many other models of communication complexity with varying levels of power. Some of these are summarised in Appendix E.

4.1 Open Problems

Alas, for all the successes we have seen in one-way communication complexity, finding a total function for which there exists an advantage for quantum communication is still open. Such an advantage would ideally be exponential, as a polynomial gap could be useless in practice if sending classical bits is cheaper [BCMdW10]. But even finding a polynomial gap has proven to be a challenge.

Aaronson and LeGall have made considerable headway towards finding a subgroup which cannot be solvable in $O(\log |G|)$ bits of communication, and conjecture that such a subgroup does exist [AGRT09]. Indeed, Aaronson called the problem one of his Ten Semi-Grand Challenges[4].

While the hidden subgroup problem might show a gap for one-way communication, there is an $O(\log |G|)$ protocol in two-way communication: Bob sends g to Alice who sends back 0 if $g \in H$ and 1 otherwise. Thus we would still need a different problem in order to show even a separation for a total function between two-way randomised classical communication and one-way quantum communication.

References

[AGRT09] Scott Aaronson, Franois Le Gall, Alexander Russell, and Seiichiro Tani. The one-way communication complexity of group membership, 2009. `arxiv:0902.3175`.

[BB14] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560, Part 1:7 – 11, 2014. Theoretical Aspects of Quantum Cryptography celebrating 30 years of {BB84}.

[4] `http://www.scottaaronson.com/blog/?p=471`

[BBC+93] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Physical Review Letters*, 70:1895–1899, Mar 1993.

[BCMdW10] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Review of Modern Physics*, 82:665–698, Mar 2010.

[BCW98] Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, pages 63–68, New York, NY, USA, 1998. ACM.

[BCWdW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87:167902, Sep 2001.

[Bra01] Gilles Brassard. Quantum communication complexity (a survey), 2001. `arxiv:quant-ph/0101005`.

[BW92] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Physical Review Letters*, 69:2881–2884, Nov 1992.

[BYJK08] Ziv Bar-Yossef, T. S. Jayram, and Iordanis Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM Journal on Computing*, 38(1):366–384, 2008.

[CB97] Richard Cleve and Harry Buhrman. Substituting quantum entanglement for communication. *Phys. Rev. A*, 56:1201–1204, Aug 1997.

[CvDNT13] Richard Cleve, Wim van Dam, Michael Nielsen, and Alain Tapp. Quantum entanglement and the communication complexity of the inner product function. *Theoretical Computer Science*, 486:11 – 19, 2013. Theory of Quantum Communication Complexity and Nonlocality.

[DPP14] Deepesh Data, Manoj M. Prabhakaran, and Vinod M. Prabhakaran. *Advances in Cryptology – CRYPTO 2014: 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II*, chapter On the Communication Complexity of Secure Computation, pages 199–216. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[FR87] Peter Frankl and Vojtěch Rödl. Forbidden intersections. *Transactions of the American Mathematical Society*, 300(1):259–286, 1987.

[FY92] Matthew Franklin and Moti Yung. Communication complexity of secure computation (extended abstract). In *Proceedings of the Twenty-fourth Annual ACM Symposium on Theory of Computing*, STOC '92, pages 699–710, New York, NY, USA, 1992. ACM.

[Gav08] Dmitry Gavinsky. Classical interaction cannot replace a quantum message. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 95–102. ACM, 2008.

[GKK+07] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 516–525, New York, NY, USA, 2007. ACM.

[Gro96] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 212–219, New York, NY, USA, 1996. ACM.

[Hol73] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:3–11, 1973.

[Ker09] Iordanis Kerenidis. Quantum multiparty communication complexity and circuit lower bounds. *Mathematical Structures in Computer Science*, 19:119–132, 2 2009.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions. In *Foundations of Computer Science, 1988., 29th Annual Symposium on*, pages 68–80, Oct 1988.

[KLGR14] Iordanis Kerenidis, Mathieu Laurire, Franois Le Gall, and Mathys Rennela. Privacy in quantum communication complexity, 2014. `arxiv:quant-ph/9508009`.

[KLL+12] Iordanis Kerenidis, Sophie Laplante, Virginie Lerays, Jeremie Roland, and David Xiao. Lower bounds on information complexity via zero-communication protocols and applications. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, FOCS '12, pages 500–509, Washington, DC, USA, 2012. IEEE Computer Society.

[MNSW98] Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, 57(1):37 – 49, 1998.

[Mon11] Ashley Montanaro. A new exponential separation between quantum and classical one-way communication complexity. *Quantum Info. Comput.*, 11(7-8):574–591, July 2011.

[Mut05] S. Muthukrishnan. Data streams: Algorithms and applications. *Foundations and Trends in Theoretical Computer Science*, 1(2):117–236, 2005.

[Nay99] Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. In *Foundations of Computer Science, 1999. 40th Annual Symposium on*, pages 369–376, 1999.

[NS96] Ilan Newman and Mario Szegedy. Public vs. private coin flips in one round communication games (extended abstract). In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 561–570, New York, NY, USA, 1996. ACM.

[RK11] Oded Regev and Bo'az Klartag. Quantum one-way communication can be exponentially stronger than classical communication. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, STOC '11, pages 31–40, New York, NY, USA, 2011. ACM.

[RP95] D. Rohrlich and S. Popescu. Nonlocality as an axiom for quantum theory, 1995. `arxiv:quant-ph/9508009`.

[Sho97] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, October 1997.

[vD12] Wim van Dam. Implausible consequences of superstrong nonlocality. *Natural Computing*, 12(1):9–12, 2012.

[VY11] Elad Verbin and Wei Yu. The streaming complexity of cycle counting, sorting by reversals, and other problems. In *Proceedings of the Twenty-second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, pages 11–25. SIAM, 2011.

[Wat00] John Watrous. Succinct quantum proofs for properties of finite groups. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, FOCS '00, pages 537–, Washington, DC, USA, 2000. IEEE Computer Society.

[Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing(preliminary report). In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, STOC '79, pages 209–213, New York, NY, USA, 1979. ACM.

[Yao93] Andrew Chi-Chih Yao. Quantum circuit complexity. In *Foundations of Computer Science, 1993. Proceedings., 34th Annual Symposium on*, pages 352–361, Nov 1993.

Appendix A Nayak's Proof of Quantum Communication Complexity

Proof of Theorem 2.1. Suppose we create an encoding $x \mapsto \{q_{x,i}|\phi_{x_i}\rangle\}$ such that x is an n-bit string and $\{q_{x,i}|\phi_{x_i}\rangle\}$ is a mixed state over m qubits. We decode this message by measuring over projection operators $\{P_x\}$. The mean probability of Bob successfully decoding an x that has been selected uniformly at random is $2^{-n}\sum_{x,i} q_{x,i}\|P_x|\phi_{x,i}\rangle\|^2$. This can be bounded by making our states pure and setting $|\phi_x\rangle = \max_{|\psi_{x,i}\rangle} \|P_x|\phi_{x,i}\rangle\|^2$:

$$\frac{1}{2^n}\sum_{x,i} q_{x,i}\|P_x|\phi_{x,i}\rangle\|^2 \leq \frac{1}{2^n}\sum_{x}\|P_x|\phi_x\rangle\|^2$$

Next, we decompose P_x into an orthonormal basis $\{|\hat{e}_{x,j}\rangle\}$ and create a subspace E that spans the code words $\{|\phi_x\rangle\}$ with projection Q:

$$\frac{1}{2^n}\sum_{x}\|P_x|\phi_x\rangle\|^2 = \frac{1}{2^n}\sum_{x,j}\langle\phi_x|\hat{e}_{x,j}\rangle\langle\hat{e}_{x,j}|\phi_x\rangle \leq \frac{1}{2^n}\sum_{x}\|Q|\hat{e}_{x,j}\rangle\|^2$$

Finally, we decompose $Q = \sum_i |e_i\rangle\langle e_i|$ where $\{|e_i\rangle\}$ is an orthonormal basis spanning E. Note that since $|\phi_x\rangle$ is conveyed in at most m qubits, $|\{|e_i\rangle\}| \leq 2^m$. Hence we conclude that

$$\frac{1}{2^n}\sum_{x,j}\|Q|\hat{e}_{x,j}\rangle\|^2 \leq \frac{1}{2^n}\sum_{x,j,i}|\langle e_i|e_{x,j}\rangle|^2 \leq \frac{1}{2^n}\sum_{i}|\langle e_i|e_i\rangle|^2 \leq 2^{m-n}.$$

□

Appendix B Distributed Deutsch-Jozsa Lower Bound

Proof of Theorem 3.1. To lower bound the Distributed Deutsch-Jozsa problem, denoted DDJ, we will partition the problem space up into rectangles (a, b, c, d) such that $DDJ(x_i, y_i) = DDJ(x'_i, y'_i)$ for any pair of inputs $(x_i, y_i), (x'_i, y'_i) \in \{x_a, x_{a+1}, ..., x_b\} \times \{y_c, y_{c+1}, ..., y_d\}$. We can lower bound the number of messages to c bits if we need 2^c rectangles to cover the entire problem space.

For this problem, we will focus on the inputs (x, x), where x is an n-bit string with hamming weight $n/2$. There are $\binom{n}{n/2} \approx 2^n/\sqrt{n}$ such inputs, which means that at least one rectangle R must contain $2^{n-c}/\sqrt{n}$ of these inputs. Because $DDJ(x, x) = 0 \ \forall\ x \in \{0,1\}^n$, all (x, y)-pairs in R must satisfy either $x = y$ to output 0 or $d(x, y) \neq n/2$ to break the promise. Characteristic analysis detailed in [BCMdW10] show that the maximum number of (x, x) inputs in R is 1.99^n [FR87]. Hence we have the inequality

$$\frac{2^{n-c}}{\sqrt{n}} \leq 1.99^n$$

Solving for c shows that $c \geq \log_2(2^n/1.99^n\sqrt{n}) = \Omega(0.007n)$. □

Note that Buhrman et al. [BCMdW10] provide this as a lower bound instead of an asymptotic lower bound, but this is only correct if $n \geq 728$. Figure 2 shows a plot of $\log_2(2^n/(1.99^n * \sqrt{n}))$ and $0.007n$ for even values of n ranging from 2 to 1000. Note that the inequality $\log_2(2^n/(1.99^n * \sqrt{n})) \geq 0.0007n$ claimed in [BCMdW10] only fits if $n \geq 728$.

We can see that the inequality holds for all values of $n \geq 728$ by differentiating the two sides of the equation. The right-hand side can be easily differentiated as $d(0.0007n)/dn = 0.0007$, and the left-hand side can be differentiated as follows:

$$\frac{d}{dn}\log_2\left(\frac{2^n}{1.99^n\sqrt{n}}\right) = \frac{d}{dn}\log_2\left(\frac{2^n}{1.99^n}\right) - \frac{d}{dn}\log_2(\sqrt{n}) = \log_2\left(\frac{2}{1.99}\right) - \frac{1}{\ln(4)n}$$

As n tends towards infinity, this derivative will tend towards $log_2(2/1.99) \approx 0.00723 > 0.0007$. The point when the derivative is greater than 0.0007 is when $n \geq 112$. Thus by the time the inequality by Buhrman et al. holds, the left hand side will be always growing at a faster rate than the right and the inequality will continue to hold.

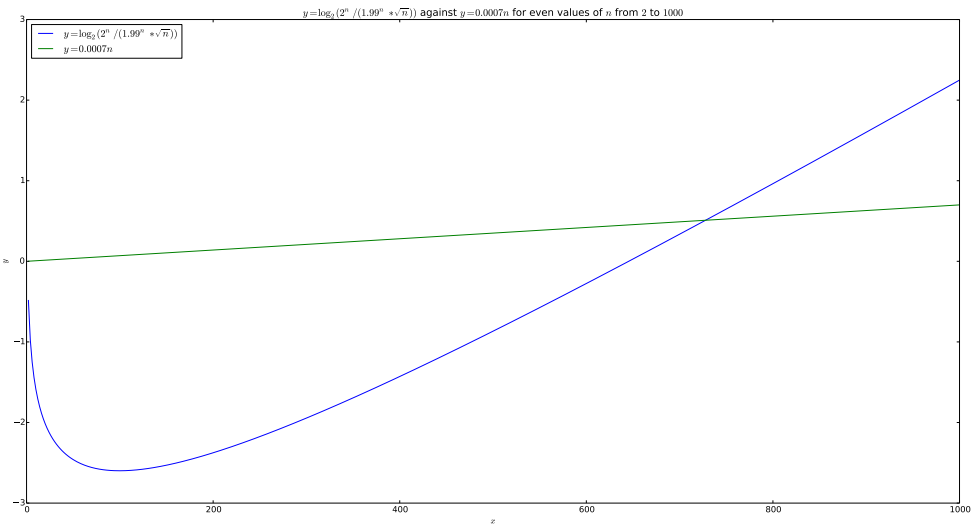

Figure 2: Plot of lower bounds of Distributed Deutsch-Jozsa for $n \in \{2, 4, ..., 998, 1000\}$

Appendix C Lower Bound for Hidden Matching

Proof of Theorem 3.3. Take any protocol for computing the Hidden Matching Problem with mean probability of failure δ across Alice and Bob's input. If we pick a certain input for Alice x, we will denote the probability of error given x as $\hat{e}_x$. By Markov's inequality, $Pr[\hat{e}_x \geq 2\delta] \leq \mathbb{E}[\hat{e}_x]/2\delta = \delta/2\delta = 1/2$, so at most half of Alice's inputs can fail more than $\delta/2$ of the time.

Assuming Alice's message to Bob is written deterministically, we define the set of Alice's inputs that map to a message τ and have error less than 2δ as S_τ. Note that $\sum_\tau S_\tau$ is equal to the number of Alice's inputs with error less than 2δ. Using Markov's inequality above, this sum must be greater than 2^{n-1}. Bar-Yossef, Jayram and Kerenidis then showed that for any message τ, the size of the set S_τ is at most $2^{\Omega(n-\sqrt{n})}$. Combining these two inequalities, the number of unique messages Alice could send to Bob is at least $2^{n-1}/2^{n-\Omega(\sqrt{n})} = 2^{\Omega(\sqrt{n})-1} = 2^{\Omega(\sqrt{n})}$. Assuming Alice's input is chosen uniformly at random, the length of each message must be at least $\Omega(\sqrt{n})$ classical bits. This bound is tight if public random bits are shared between Alice and Bob, as Alice can send $O(\sqrt{n})$ bits x_i, where i is chosen from the random string. By the birthday paradox, the expected probability that a pair of indices were in Bob's matching is 1. □

Appendix D Permutation Invariance

Permutation Invariance is a problem first studied in the context of communication complexity by Montanaro [Mon11], and has connections both to the Hidden Subgroup Membership problem discussed in Section 3.2 and the Hidden Matching problem discussed in Section 3.3.

Let M be an $n \times n$ permutation and x be an n-bit string. We define the n-bit string $Mx = x_{M_i}$ for $i \in \{0, ..., n-1\}$. We also define $d(a, b) = \sum_{i \in n}(a_i \oplus b_i)$ as the Hamming distance between a and b. The problem and its quantum protocol are explained below for a parameter β:

Protocol PERMUTATION-INVARIANCE

ALICE'S INPUT: $x \in \{0,1\}^n$.

BOB'S INPUT: an $n \times n$ permutation M.

PROMISE: $Mx = x$ or $d(Mx, x) \geq \beta|x|$.

BOB'S OUTPUT: 0 if $Mx = x$, 1 otherwise.

1 Alice sends the state $|\psi_x\rangle = \frac{1}{\sqrt{x}} \sum_{i,x_i=1} |i\rangle$ to Bob.
2 Bob adds an ancilla qubit in the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.
3 Bob applies the operation $M = \sum_{i \in \{0,\ldots,n-1\}} |M_i\rangle\langle i|$ controlled on the ancilla qubit.
4 Bob applies a Hadamard to the ancilla qubit and measures in the computational basis.
5 If Bob measures 0 they output 0, otherwise they output 1.

This protocol can be completed using $O(\log n)$ qubits of communication.

The state after step 2 of the operation is $(|0\rangle|\psi_x\rangle + |1\rangle|\psi_x\rangle)/\sqrt{2}$. After applying the controlled M gate, this becomes $(|0\rangle|\psi_x\rangle + |1\rangle|\psi_{Mx}\rangle)/\sqrt{2}$, where $|\psi_{Mx}\rangle = \sum_{i,x_i=1} |M_i\rangle$. If $Mx = x$ then $|\psi_{Mx}\rangle = |\psi_x\rangle$, so the state is not changed and Bob measures 0 with absolute certainty.

If $Mx \neq x$ on the other hand, then the state after step 4 becomes

$$\frac{1}{2}(|0\rangle(|\psi_x\rangle + |\psi_{Mx}\rangle) + |1\rangle(|\psi_x\rangle - |\psi_{Mx}\rangle)).$$

The probability of measuring 0 in this case is $\frac{1}{4}||\psi_x\rangle + |M\psi_x\rangle\|_2^2 = \frac{1}{2}(1 + \langle\psi_{Mx}|\psi_x\rangle)$[5]. Since $d(Mx, x) \geq \beta|x|$, we can see that $\langle\psi_{Mx}|\psi_x\rangle \leq 1 - \beta$, so this protocol fails with probability $1 - \beta/2$.

This problem can be seen as a generalisation of subgroup membership: Alice's input is a $|G|$-bit string x where $x_i = 1$ iff $i \in H$ for all $i \in G$, and Bob's permutation M_g maps elements $i \in G$ to gi. If $g \in H$, then $M_g x = x$, otherwise $d(M_g x, x) = 2|x|$. So already there is potentially a quadratic gap between quantum and classical communication for this problem, depending on whether or not such a gap exists for subgroup membership.

The problem also relates to Hidden Matching if we only consider permutations which are completely made of disjoint pairs – in other words, the permutation pairs bits up and swaps those pairs. Bob can build a perfect matching M' by adding edges (i, j) if the permutation matrix M swaps i with j. Hence this problem can be reduced to Hidden Matching, with the difference being that Bob no longer has w as an input.

How well can we perform classically? Montanaro [Mon11] provided a lower bound by looking at a subproblem, where Alice has a $2n$-bit string x such that $|x| = n$, Bob has a $2n \times 2n$ permutation matrix M and either $Mx = x$ or $d(Mx, x) \geq n/8$. Montanaro's proof breaks the problem up into two probability distributions: $\mathcal{D}_0$ when $d(Mx, x) \geq n/8$ and $\mathcal{D}_1$ when $Mx = x$. Their conclusion was that if Alice sends fewer than $\frac{n^{7/16}}{e\ln 2} - O(\log n)$ classical bits, then $\mathcal{D}_0$ and $\mathcal{D}_1$ becomes indistinguishable from the normal distribution within constant factors, and thus the distributions are indistinguishable from each other. Thus we have the following:

Theorem D.1. *Any randomised one-way protocol for Permutation Invariance requires $O(n^{7/16})$ classical bits of communication.*

Appendix E Other Models

Throughout this report we have focused on one-way communication complexity. But messaging between two parts is not the only model for distributed computation, and there are a number of other communication models which have also been studied. Some of these models are briefly summarised below.

E.1 Zero Communication Complexity

Also known as 'spooky communication complexity' [Bra01], this is the study of distributed computation problems which can be solved using only entangled qubits, without any communication whatsoever. Complexity is measured in the number of entangled pairs that have been used during computation.

[5] Since our amplitudes are real numbers, $\langle\psi_{Mx}|\psi_x\rangle = \langle\psi_x|\psi_{Mx}\rangle$.

Much like how Holevo's theorem limits strengths of quantum communication, Alice and Bob are limited by the non-signalling properties of entanglement.

Zero communication complexity is strongly related to non-locality [BCMdW10]. There have been a number of quantum communication problems shown to be solvable in this model of computation as well, including both the Distributed Deutsch-Jozsa problem and the Hidden Matching problem. The protocols have also been used to prove lower bounds on information complexity: If there is a protocol for computing a function with information complexity I, there is also a zero communication protocol for computing the same function such that the function has probability greater than $2^{-O(I)}$ of not aborting, and the protocol has high probability of succeeding on condition of it not aborting [KLL+12].

E.2 Simultaneous Message Passing

In the Simultaneous Message Passing model, Alice and Bob are given inputs x and y, respectively, and send messages one-way to a third player called the Referee without any communication between each other. The Referee must then output $f(x, y)$ for some function f. Another way of picturing this model is if Alice and Bob sent a single message to each other at exactly the same time, so that one party could not have received the other parties message before sending theirs, and then for both players to output the same result $f(x, y)$.

It is already established that total functions are exponentially separated in this model: Newman and Szegedy [NS96] proved that equality of two n-bit strings in this model required $O(\sqrt{n})$ bits of communication. On the other hand, Buhrman et al. [BCWdW01] showed that the same problem in this model could be solved in $O(\log n)$ qubits of communication. It was also shown by Gavinsky et al. [GKK+07] that α-Matching can be solved in the Simultaneous Message Passing model using $O(\log n)$ qubits of communication.

E.3 Non-Local Boxes

Non-local boxes act as a similar resource to entanglement, where Alice and Bob are each given a black box. If Alice inputs a bit x into her box, and Bob inputs a bit y into his box, then Alice will get bit a and Bob will get bit b, subject to the rule that $x \wedge y = a \oplus b$. They were first proposed by Rohrlich and Popescu [RP95] to see what a maximally non-local experiment would look like. Since then, several properties have been proven and summarised by Buhrman et al. [BCMdW10]. Of particular note, it was shown by van Dam [vD12] that the existence of non-local boxes would mean that all distributed computation could be performed using a single bit of communication between parties.