One-Way Quantum Communication Complexity

Dominic Moylett*

Quantum Engineering Centre for Doctoral Training,
University of Bristol

March 18, 2016

Abstract

Communication complexity is the study of how much information two or more parties need to share with each other in order to perform joint computation of a problem. There are many benefits to this model of computation, from many lower bound proofs to applications in cryptography and data streaming. In this report, we will summarise the most exciting recent trends in the one-way form of quantum communication complexity, look at what quantum advantages already exist and what problems are still open.

1 Introduction

2 Communication Complexity

Communication complexity was developed by Yao in 1979 as an analysis of distributed computing [Yao79]. Under this model of computation, 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 protocol established before computation, 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 the two of them to compute a solution. We offer a sketch of a typical communication complexity problem in Figure 1.

A naïve solution to any communication complexity problem can be completed using n bits of communication: Alice sends x to Bob and Bob then does all of the computation for $f(x, y)$. So the question then becomes if it is possible to compute $f(x, y)$ with $o(n)$ bits of communication.

In quantum communication complexity, first proposed in 1993 by Yao [Yao93], we also have parties Alice and Bob given inputs x and y, respectively, and have to produce $f(x, y)$ 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 Burhman and Cleve [CB97] does not use the physical sending of data between the parties, but instead uses the measurement of entangled particles for remote computation. Our focus for this report however will be on the former. Note that 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. Likewise, it is possible to simulate communicating l qubits by use of l entangled pairs and $2l$ classical bits of communication, via quantum teleportation [BBC$^+$93].

2.1 One-Way Communication Complexity

A specific subset of communication complexity, which we will focus on throughout this report, is one-way communication complexity. In this model of computation, Alice and Bob again receive respective inputs

*dominic.moylett@bristol.ac.uk

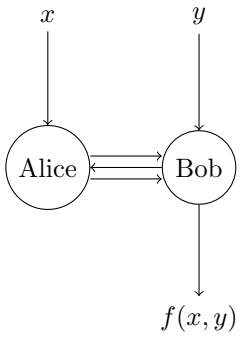

Figure 1: An example of communication complexity. Alice and Bob begin computation with bit strings x and y. They then exchange a series of messages based on a pre-determined protocol. Computation completes when Bob outputs $f(x, y)$.

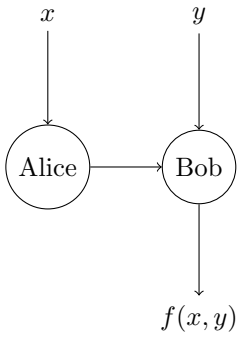

Figure 2: An example of one-way communication complexity. Alice receives input x and Bob receives input y. Alice then sends a single message to Bob. Using this additional information from Alice, Bob performs and outputs the computation $f(x, y)$.

$x, y \in \{0, 1\}^n$ with the aim of Bob being able 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. This is illustrated in 2.

Like general communication complexity, one-way communication complexity can also be solved naïvely in n bits of communication by Alice sending x to Bob. So once again, our question is if we can do better with less communication.

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. In particular, a consequence of Holevo's theorem [Hol73], meant that while a message of n bits could be compressed into a message of at most $n - 1$ qubits, the same message could not be retrieved from said qubits. We phrase this as a one-way communication complexity problem below.

Problem COMMUNICATION
ALICE'S INPUT: $x \in \{0, 1\}^n$.
BOB'S INPUT: none.
BOB'S OUTPUT: x.

By Holevo's theorem, we know that the above problem requires Alice to send $\Omega(n)$ qubits to Bob. A later proof by Nayak [Nay99] shows that if we compress a message of n bits into a message of $m < n$ qubits, the probability that we can successfully retrieve all n bits again is at most 2^{m-n}. The only exception to this is when Alice and Bob have shared entangled qubits between each other prior to computation, in which case superdense coding can be used to send two bits by sending one qubit [BW92]. Brassard [Bra01] explains that if we have unlimited entanglement between Alice and Bob then we can only at

most reduce the amount of communication by a factor of 1/2, as quantum teleportation can be used to send qubits from one party to the other using entangled pairs and communicating with classical bits [BBC+93].

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, studied in [CvDNT13]:

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. 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. A summary of this proof is given in [BCMdW10].

While it is clear that we cannot solve all problems in fewer qubits than we can in classical bits, we will in the next two sections look at more specific problems which we can solve by communication of fewer qubits.

3 Functions

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. The problem, and its quantum protocol as described in Buhrman et al. [BCMdW10], are stated as follows, where $\mathcal{D}(x,y)$ 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 $\mathcal{D}(x,y) = n/2$.
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 $\mathcal{D}(x,y) = n/2$.

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? It was shown in Appendix 4 of Buhrman et al. [BCMdW10] that any deterministic protocol must send at least $0.007n$ bits from Alice to Bob, otherwise the protocol will fail for some input. However, this quantum advantage is lost when compared to randomised classical computation for some positive integer c:

Protocol DISTRIBUTED DEUTSCH-JOZSA

1 Alice picks c unique items $a \in \{0,1\}^n$ uniformly at random.
2 Alice sends $(a_i, x_{a_i})\ \forall i \in \{0, ..., c-1\}$.
3 Bob returns 0 if $x_{a_i} = y_{a_i}\ \forall i \in \{0, ..., c-1\}$, and 1 otherwise.

Bob always returns 0 if $x = y$, so our protocol has no true negative results. If $\mathcal{D}(x, y) = n/2$, then the protocol fails if $x_{a_i} = y_{a_i} \forall i \in \{0, ..., c-1\}$, which happens with probability $1/2^c$. Finally, note that each a_i can be communicated in $\log n$ bits, so the overall communication required is $c(\log n + 1)$ bits.

3.2 Subgroup Membership

One disadvantage of the Distributed Deutsch-Jozsa problem that we saw above is that a quantum protocol has no advantage over a randomised classical protocol. Another disadvantage is that the problem is a *partial function*, in that it requires a *promise* on the input: that either the two bit strings are identical or they differ in exactly $n/2$ places. This promise results in a very contrived problem with little practical application. This begs the question: Are there any functions which do not rely on promises that yield a quantum advantage under the communication complexity model – better known as *total functions*?

Sadly, no significant gap has been shown for total functions in the one-way model, even though two-way communication has shown polynomial gaps [BCW98], and other models such as Simultaneous Message Passing have shown exponential gaps [BCWdW01]. But there are some problems which have shown potential. Aaronson et al. [AGRT09] were the first to explore the problem of subgroup membership as potentially yielding a quantum advantage in the one-way comunication complexity setting. Their protocol for solving subgroup membership, based on an earlier quantum 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 $|yH\rangle = \frac{1}{\sqrt{|H|}} \sum_{yh \in H} |h\rangle$. If $y \in H$ then $yh \in H\ \forall h \in H$, so $|yH\rangle = |H\rangle$. Therefore, 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

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

Measuring the final 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 get a negligible error probability.

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 $\Omega(\log |G|)$ bits of communication. Aaronson et al. [AGRT09] have conjectured that this bound still holds even if we relax soundness to $1/3$ and tighten completeness to $2/3$.

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 polynomial gap between classical and quantum communication for a total function. The problem is that there are many subgroup membership problems which *can* be solved classically within these constraints. One example is that of normal subgroups.

Aaronson et al. [AGRT09] showed that this problem can be solved with error probability of 1/2 using only $O(\log |G|)$ bits of communication.

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 $\mathrm{Char}(g)$. Finally, if H is a normal subgroup of G, then let

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

Finally, 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.

If $y \in H$ then $y \in \ker(\mu)$, so $\mu(y) = \mu(1)$ and thus the protocol will always correctly report a member of H. So now we only need to consider the case where the $y \notin H$ yet the protocol accepts. This will occur if Alice picks a character μ such that $y \in \ker(\mu)$. We shall call the set of all characters that would cause an error B. Based on Alice's selection criteria, 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 $\forall\ h \in H, h \in K$ and $hyh^{-1} \in K$, but $hyh^{-1} \notin H$.

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 $H \subseteq \ker(\chi)$, and also that $y \in \ker(\chi)$ because $\chi \in B$. We can conclude 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

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

Thus this protocol fails with probability at most 1/2. Finally, we note that is is possible to describe a character μ sent from Alice to Bob in $\lceil \log |G| \rceil$ bits, as $\mu \in \Lambda_H$ and $|\lambda_H| \leq |Char(G)| \leq |G|$, so this protocol only needs $O(\log |G|)$ classical bits of communication. So in order to find a subgroup membership problem which provides an exponential gap between random and quantum computation, we need to consider non-normal subgroups.

3.3 Permutation Invariance

4 Relations

4.1 Hidden Matching

A matching M is a sequence of 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 inner product of any two different states in the basis in step 2 of the above protocol is

$$\langle \psi_{i\pm j} | \psi_{i'\pm j'} \rangle = \frac{\langle i|i'\rangle \pm \langle i|j'\rangle \pm \langle j|i'\rangle \pm \langle j|j'\rangle}{2} = 0$$

The only exception is comparing the states $|\psi_{i+j}\rangle$ and $|\psi_{i-j}\rangle$, where:

$$\langle \psi_{i+j} | \psi_{i-j} \rangle = \frac{\langle i|i\rangle - \langle j|j\rangle}{2} = 0$$

Hence this basis is orthogonal and can be measured. The probability of measuring $|\psi_{i\pm j}\rangle$ is

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

.

Note that 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 only measure the state $|\psi_{i+j}\rangle$. Likewise, if $x_i \oplus x_j = 1$ then $P(i+j) = \frac{1}{2n}(1-1) = 0$, so Bob will only measure the state $|\psi_{i-j}\rangle$. Thus this quantum protocol will always succeed at finding a matching pair, though the specific pair found occurs with probability $\frac{1}{2n}$.

As for classical one-way protocols, it was shown in [BYJK08] that a random one-way protocol needs $\Omega(\sqrt{n})$ bits of communication, thus providing an exponential gap. Furthermore, this bound is known to be 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.

So have we now found an example of a total function that separates quantum and classical one-way communication complexity? Alas, we have not. This is because a function has a one-to-one mapping from inputs to outputs. But 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 $\frac{n}{2}$ acceptable outputs for this problem. This makes the hidden matching problem not a function, but a *relation*.

4.2 α-Matching

α-matching is a promise function related to the hidden matching problem described earlier. It was first studied for one-way communication complexity by Gavinsky et al. [GKK$^+$07]. In α-matching, the matching given to Bob consists of αn pairs for, such that no index occurs in more than one pair. 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 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 a perfect matching M'.
2 Alice and Bob engage in the protocol for Hidden Matching on M' to get tuple $(i, j, x_i \oplus x_j)$
3 If $(i,j) \notin M$ then Bob aborts.
4 Bob outputs 0 if $|\psi_{i\pm j}\rangle = \frac{\pm 1}{\sqrt{2}}(|i\rangle + |j\rangle)$ and 1 if $|\psi_{i\pm j}\rangle = \frac{\pm 1}{\sqrt{2}}(|i\rangle - |j\rangle)$.

First note that Bob only needs one bit of Mx to figure out which of the two cases holds. This is because of the promise we are given: either $w_i = (Mx)_i \; \forall \; i \in \{0, ..., \alpha n - 1\}$, or $w_i \neq (Mx)_i \; \forall \; i \in \{0, ..., \alpha n - 1\}$. So if we can find one pair (x_i, x_j) such that $(i,j) \in M$, then we have solved this problem. Hence one can think of this as a reduction to the Hidden Matching problem described in Section 4.1.

If $\alpha = 1/2$, then $M = M'$ and thus we have the same protocol as we had for Hidden Matching, so the protocol will always succeed in that case. Indeed, 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] used Fourier analysis to argue 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 described later.

If Alice's message is deterministic based on x, then the c bits Alice sends allow Bob to define a set $A \subseteq \{0,1\}^n$ such that $|A| \geq 2^{n-c}$ of bit strings from which they must be able to guess Alice's input x. The probability of Bob succeeding then is the probability of them picking a bit string $x' \in A$ such that $Mx = Mx'$. We can define the probability of Bob finding a 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. Gavinsky et al. show that the Fourier transform of our probability distribution is related to the indicator function of our set of possible Alice messages A:

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

Note that there is exactly one string $s \in \{0,1\}^{\alpha n}$ such that $My = s\ \forall\ y \in A$. Because we are summing over all αn-bit strings, we can remove that summation:

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

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

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

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

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

Now we need to consider what this probability over M is. Suppose the Hamming weight of v is $|v| = k$ for even k[1]. Lemma 5 of [GKK+07] showed that the fraction of partial matchings in $\mathcal{M}_\alpha$ that

[1]The Hamming weight of M^Ts must be even and at most $2\alpha n$, so this restriction is acceptable for the problem at hand.

satisfy $M^T s = v$ is equal to $\binom{\alpha n}{k/2}/\binom{n}{k}$. In other words, the number of α-matchings which satisfy the condition 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. We can now re-write this probability as a sum over even values of 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] to find that

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

Noting that $1+\delta > 1$, $|A| \geq 2^{n-c}$, $c \leq \gamma\epsilon\sqrt{n/\alpha}$ and estimating that $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$ we can bound this even further:

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

A small value for γ will ensure that this sum converges to at most $\epsilon^2/2$.

Now that the lower part of the sum has been bound, we 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:

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

We can use this and our previous estimate for binomial coefficients to bound this equation:

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

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

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

Again, by choosing γ to be a small enough constant, this equation can be bounded by $\epsilon^2/2$. By adding these two cases together, we find that the squared total variation distance is at most ϵ^2. To conclude this proof, we use Jensen's inequality to show that

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

Square-rooting these terms gives us that the total variation distance between Bob's probability distribution and the uniform becomes arbitrarily small if Alice sends much less than $\sqrt{n/\alpha}$ bits of classical data to Bob. Thus an exponential separation is shown.

5 Conclusion

5.1 Open Problems

5.2 Other Areas

References

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

[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. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[BCMdW10] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Rev. Mod. Phys.*, 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. *Phys. Rev. Lett.*, 87:167902, Sep 2001.

[Bra01] Gilles Brassard. Quantum communication complexity (a survey), 2001.

[BW92] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 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, 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.

[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.

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

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

[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.

[Nay99] A. 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.

[Wat00] J. 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] A. Chi-Chih Yao. Quantum circuit complexity. In *Foundations of Computer Science, 1993. Proceedings., 34th Annual Symposium on*, pages 352–361, Nov 1993.