Exponential Separation of Quantum and Classical One-Way Communication Complexity

Ziv Bar-Yossef
IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120, USA.
ziv@almaden.ibm.com

T. S. Jayram
IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120, USA.
jayram@almaden.ibm.com

Iordanis Kerenidis*
Computer Science Division
U.C. Berkeley
Berkeley, CA 94720, USA.
jkeren@cs.berkeley.edu

ABSTRACT

We give the first exponential separation between quantum and bounded-error randomized one-way communication complexity. Specifically, we define the Hidden Matching Problem HM_n: Alice gets as input a string $\mathbf{x} \in \{0,1\}^n$ and Bob gets a perfect matching M on the n coordinates. Bob's goal is to output a tuple $\langle i, j, b \rangle$ such that the edge (i, j) belongs to the matching M and $b = x_i \oplus x_j$. We prove that the quantum one-way communication complexity of HM_n is $O(\log n)$, yet any randomized one-way protocol with bounded error must use $\Omega(\sqrt{n})$ bits of communication. No asymptotic gap for one-way communication was previously known. Our bounds also hold in the model of Simultaneous Messages (SM) and hence we provide the first exponential separation between quantum SM and randomized SM with public coins.

For a Boolean decision version of HM_n, we show that the quantum one-way communication complexity remains $O(\log n)$ and that the 0-error randomized one-way communication complexity is $\Omega(n)$. We prove that any randomized *linear* one-way protocol with bounded error for this problem requires $\Omega(\sqrt[3]{n \log n})$ bits of communication.

Categories and Subject Descriptors

F.1.1 [**Theory of Computation**]: Computation by Abstract Devices—*Models of Computation*; E.4 [**Data**]: Coding and Information Theory

General Terms

Theory

*Part of this work was done while the author was visiting IBM Almaden Research Center.

Keywords

Communication complexity, quantum computation, separation, hidden matching

1. INTRODUCTION

The investigation of the strength and limitations of quantum computing has become an important field of study in theoretical computer science. The celebrated algorithm of Shor [18] for factoring numbers in polynomial time on a quantum computer gives strong evidence that quantum computers are more powerful than classical ones. The further study of the relationship between quantum and classical computing in models like black-box computation, communication complexity, and interactive proof systems help towards a better understanding of quantum and classical computing.

In this paper we answer an open question about the relative power of quantum *one-way communication protocols*. We describe a problem which can be solved by a quantum one-way communication protocol exponentially faster than any classical one. No asymptotic gap was previously known. We prove a similar result in the model of Simultaneous Messages.

Communication complexity is a central model of computation with numerous applications. It has been used for proving lower bounds in many areas including Boolean circuits, time-space tradeoffs, data structures, automata, formulae size, etc. Examples of these applications can be found in the textbook of Kushilevitz and Nisan [12]. A communication complexity problem is defined by three sets X, Y, Z and a relation $\mathcal{R} \subseteq X \times Y \times Z$. The two players, Alice and Bob, are given inputs $x \in X$ and $y \in Y$ respectively. Their goal is to output an answer $z \in Z$, such that $(x, y, z) \in \mathcal{R}$. The communication complexity of the problem is the number of bits Alice and Bob must exchange in the best protocol that outputs such an answer z, for the worst case inputs x, y. The two players have unlimited computational power. The model of communication complexity for functions was introduced by Yao [19] and was generalized to relations by Karchmer and Wigderson [8]. One important special case of the above model is *one-way communication complexity*, where Alice is allowed to send a single message to Bob, after which Bob computes the output. *Simultaneous Messages (SM)* is a variant in which Alice and Bob cannot communicate directly with each other; instead, each of them sends a single message to a third party, the "referee", who announces the output based on the two messages.

Depending on the kind of allowed protocols we can define different measures of communication complexity for a problem $\mathcal{P}$. The classical *deterministic* communication complexity of $\mathcal{P}$ is the number of bits exchanged during the execution of the optimal deterministic protocol for $\mathcal{P}$. In a *bounded-error randomized* protocol with error probability $\delta > 0$, both players have access to *public* random coins and the output of the protocol is required to be correct with probability at least $1 - \delta$. The cost of such a protocol is the number of bits Alice and Bob exchange on the worst-case choice of inputs and of values for the random coins. The randomized communication complexity of $\mathcal{P}$ (w.r.t. δ) is the cost of the optimal randomized protocol for $\mathcal{P}$. In a *0-error randomized* protocol (a.k.a. Las Vegas protocol) the players need to output the correct value with probability 1. The cost of such a protocol is the *expected* number of bits Alice and Bob exchange on the worst-case choice of inputs. These complexity measures can also be specialized by restricting the communication model to be SM or one-way communication. An interesing variant for randomized protocols in the SM model is when the random coins are restricted to be *private*[1].

The definition of *quantum* communication complexity is due to Yao [21]. Similarly to the classical case, quantum communication complexity, apart from being of interest in itself, has been used to prove bounds on quantum formulae size, automata, data structures, etc. (e.g., [21, 10, 17]). In this setting, Alice and Bob hold qubits, some of which are initialized to the input. In a communication round, each player can perform some arbitrary unitary operation on his/her part of the qubits and send some of them to the other player. At the end of the protocol they perform a measurement and decide on an outcome. The output of the protocol is required to be correct with probability $1 - \delta$, for some $\delta > 0$. The quantum communication complexity of $\mathcal{P}$ is the number of qubits exchanged in the optimal bounded-error quantum protocol for $\mathcal{P}$. It can be shown that the quantum communication is as powerful as bounded-error randomized communication with private coins[2], even when restricted to variants such as one-way communication and SM. It is a natural and important question to ask whether quantum channels can significantly reduce the amount of communication necessary to solve certain problems.

It is known that randomized one-way communication protocols can be much more efficient than deterministic protocols. For example, the equality function can be solved by a $O(1)$ randomized one-way protocol, though its deterministic one-way communication complexity is $\Omega(n)$ (cf. [12]). However, the question of whether quantum one-way communication could be exponentially more efficient than the randomized one remained open. We resolve this in the affirmative, by exhibiting a problem for which the quantum complexity is exponentially smaller than the randomized one.

1.1 Related work

The area of quantum communication complexity was introduced by Yao [21]. Since then, a series of papers have investigated the power and limitations of quantum communication complexity. Buhrman, Cleve, and Wigderson [5] described a relation $\mathcal{R}$ with deterministic communication complexity of $\Theta(n)$ and 0-error quantum communication complexity of $\Theta(\log n)$. However, the bounded-error randomized communication complexity of this problem is $O(1)$. An exponential separation with respect to bounded-error randomized protocols was given by Ambainis *et al.* [2] in the so called sampling model. However, the separation does not hold in the presence of public coins. Buhrman *et al.* [4] were able to solve the equality problem in the SM model with a quantum protocol of complexity $O(\log n)$ rather than the $\Theta(\sqrt{n})$ bits necessary in any bounded-error randomized SM protocol with private coins [14, 3]. Again, if we allow the players to share random coins, then equality can be solved classically with $O(1)$ communication.

Ran Raz [16] was the first to show an exponential gap between the quantum and the bounded-error public-coin randomized communication complexity models. He described a relation $\mathcal{P}_1$ with an efficient quantum protocol of complexity $O(\log n)$. He then proved a lower bound of $\Omega(n^{1/4})$ on the classical randomized communication complexity of $\mathcal{P}_1$. Since the quantum protocol given for $\mathcal{P}_1$ uses two rounds, the separation holds only for protocols that use two rounds or more. The definition of $\mathcal{P}_1$ was motivated, in part, by another relation $\mathcal{P}_0$. The latter was first introduced by Kremer [11] who showed that $\mathcal{P}_0$ is a complete problem for quantum one-way communication complexity (in particular, it has a $O(\log n)$ quantum one-way protocol). However, no lower bound is given for the one-way randomized communication complexity of $\mathcal{P}_0$. Proving an exponential separation of classical and quantum one-way communication complexity has been an open question since.

Klauck [10] proved that the 0-error quantum one-way communication complexity of *total* functions (i.e., problems $\mathcal{R} \subseteq X \times Y \times Z$, for which *every* $x \in X$ and $y \in Y$ have exactly *one* $z \in Z$ with $(x, y, z) \in \mathcal{R}$) is equal to the classical deterministic one. It is still an open question whether for total functions quantum and bounded-error randomized one-way communication complexity are polynomially related.

Subsequent to our work, Aaronson [1] showed that for any Boolean function f, the deterministic one-way communication complexity of f is $O(\log |Y| \cdot Q^1(f) \cdot \log Q^1(f))$, where $Q^1(f)$ is the bounded-error quantum one-way communication complexity of f; namely, if the given communication problem is a Boolean function in which Bob's domain is small, then deterministic one-way communication complexity is almost as efficient as bounded-error quantum one-way communication complexity.

1.2 Our results

Our main result is the definition and analysis of the communication complexity of the Hidden Matching Problem. This provides the first exponential separation between quantum and classical one-way communication complexity.

The Hidden Matching Problem:

Let n be a positive even integer. In the Hidden Matching Problem, denoted HM_n, Alice is given $\mathbf{x} \in \{0,1\}^n$ and Bob is given $M \in \mathcal{M}_n$ ($\mathcal{M}_n$ denotes the family of all possible perfect matchings on n nodes). Their goal is to output a tuple $\langle i, j, b \rangle$ such that the edge (i, j) belongs to the matching M and $b = x_i \oplus x_j$.

[1]The difference in complexity between public and private coins for the other models is only $O(\log n)$ [13].

[2]As noted earlier, the distinction between public and private coins is significant only for the SM model.

This problem is new and we believe that its definition plays the major role in obtaining our result. The inspiration comes from the recent work by Kerenidis and de Wolf on Locally Decodable Codes [9]. Let us give the intuition why this problem is hard for communication complexity protocols. Suppose (to make the problem even easier) that Bob's matching M is restricted to be one of n fixed disjoint matchings on $\mathbf{x}$. Bob's goal is to find the value of $x_i \oplus x_j$ for some $(i, j) \in M$. However, since Alice has no information about which matching Bob has, her message needs to contain information about the parity of at least one pair from each matching. Hence, she needs to communicate parities of $\Omega(n)$ different pairs to Bob. It can be shown that such message must be of size $\Omega(\sqrt{n})$. In Section 5 we turn this intuition into a proof for the randomized one-way communication complexity of HM_n. We also show that our lower bound is tight by describing a randomized one-way protocol with communication $O(\sqrt{n})$. In this protocol, Alice just sends $O(\sqrt{n})$ random bits of her input. By the birthday paradox, with high probability, Bob can recover the value of at least one of his matching pairs from Alice's message.

Remarkably, this problem remains easy for quantum one-way communication. Alice only needs to send a uniform superposition of her string $\mathbf{x}$, hence communicating only $\log n$ qubits. Bob can perform a measurement on this superposition which depends on the matching M and then output the parity of some pair in M. In Section 4 we describe the quantum protocol in more detail.

In section 6 we show that HM_n also provides the first exponential separation between quantum SM and randomized SM with public coins. Previously such a bound was known only in the private coins model.

Our main result exhibits a separation between quantum and classical one-way communication complexity for a relation. Ideally, one would like to prove such a separation for the most basic type of problems—total Boolean functions. The best known separation between quantum and classical communication complexity (even for an arbitrary number of rounds) for such functions is only quadratic [5]. In fact, it is very conceivable that for total functions, the two models are polynomially related. Raz's result [16] shows an exponential gap for a partial Boolean function (i.e., a Boolean function that is defined only on a subset of the domain $\mathcal{X} \times \mathcal{Y}$) and for two-way communication protocols.

We consider a partial Boolean function induced by the Hidden Matching Problem, defined below. In the definition we view each matching $M \in \mathcal{M}_n$ as an $\frac{n}{2} \times n$ edge-vertex incidence matrix. For two Boolean vectors $\mathbf{v}, \mathbf{w}$, we denote by $\mathbf{v} \oplus \mathbf{w}$ the vector obtained by xoring $\mathbf{v}$ and $\mathbf{w}$ coordinatewise. For a bit $b \in \{0, 1\}$, we denote by $\mathbf{b}$ the vector all of whose entries are b.

The Boolean Hidden Matching Problem:

Let n be a positive even integer. In the Boolean Hidden Matching Problem, denoted BHM_n, Alice is given $\mathbf{x} \in \{0, 1\}^n$ and Bob is given $M \in \mathcal{M}_n$ and $\mathbf{w} \in \{0, 1\}^{n/2}$, which satisfy the following promise: either $M\mathbf{x} \oplus \mathbf{w} = \mathbf{1}$ (a Yes instance) or $M\mathbf{x} \oplus \mathbf{w} = \mathbf{0}$ (a No instance).

The same quantum protocol that solves HM_n also solves BHM_n with $O(\log n)$ qubits. We were unable to extend the randomized lower bound for HM_n to a similar lower bound for BHM_n. Yet, we believe that BHM_n should also exhibit an exponential gap in its quantum and classical one-way communication complexity. We give a strong indication of that with two lower bounds. First, we prove an $\Omega(n)$ lower bound on the 0-error randomized one-way communication complexity of BHM_n. We then show that a natural class of randomized bounded-error protocols require $\tilde{\Omega}(\sqrt[3]{n})$ bits of communication to compute BHM_n. The protocols we refer to are *linear*; that is, Alice and Bob use the public coins to choose a random matrix $\mathbf{A}$, and Alice's message on input $\mathbf{x}$ is simply $\mathbf{Ax}$. These protocols are natural for our problem, because what Bob needs to compute is a linear transformation of Alice's input. In particular, the $O(\sqrt{n})$ communication protocol that we described earlier is trivially a linear protocol. Generalizing this lower bound to the case of non-linear randomized protocols remains an open problem. These results are described in Section 7.

2. PRELIMINARIES

2.1 Information theory

Throughout the paper we use basic notions and facts from information theory, which we briefly review next. We refer the reader to the textbook of Cover and Thomas [6] for details and proofs.

We deal only with finite discrete probability spaces. The distribution of a random variable X is denoted by μ_X, and let $\mu_X(x) \stackrel{\mathrm{def}}{=} \Pr[X = x]$. The *entropy* of X (or, equivalently, of μ_X) is $\mathrm{H}(X) \stackrel{\mathrm{def}}{=} \sum_{x \in \mathcal{X}} \mu_X(x) \log \frac{1}{\mu_X(x)}$, where $\mathcal{X}$ is the domain of X. The entropy of a Bernoulli random variable with probability of success p is called the *binary entropy function* of p and is denoted $H_2(p)$. The *joint entropy* of X and Y is the entropy of the joint distribution μ_{XY} of X and Y. The *conditional entropy* of X given an event A, denoted $\mathrm{H}(X|A)$, is the entropy of the conditional distribution of μ_X given A. The *conditional entropy* of X given Y is $\mathrm{H}(X|Y) \stackrel{\mathrm{def}}{=} \sum_{y \in \mathcal{Y}} \mu_Y(y)\, \mathrm{H}(X|Y = y)$, where $\mathcal{Y}$ is the domain of Y. The *mutual information* between X and Y is $\mathrm{I}(X; Y) \stackrel{\mathrm{def}}{=} \mathrm{H}(X) - \mathrm{H}(X|Y) = \mathrm{H}(Y) - \mathrm{H}(Y|X)$. The *conditional mutual information* between X and Y given Z is $\mathrm{I}(X; Y|Z) = \mathrm{H}(X|Z) - \mathrm{H}(X|Y, Z) = \mathrm{H}(Y|Z) - \mathrm{H}(Y|X, Z)$.

Some basic properties of entropy and mutual information we are using in this paper are the following.

Theorem 1. *Let X, Y, Z be random variables.*

1. $\mathrm{H}(X) \le \log |\mathcal{X}|$, *where $\mathcal{X}$ is the domain of X. Equality holds iff X is uniform on $\mathcal{X}$.*
2. *Conditioning reduces entropy:* $\mathrm{H}(X|Y) \le \mathrm{H}(X)$. *Equality holds iff X, Y are independent.*
3. *Data processing inequality: For any function f,* $\mathrm{I}(X; f(Y)) \le \mathrm{I}(X; Y)$.
4. *Chain rule for mutual information:* $\mathrm{I}(X; Y, Z) = \mathrm{I}(X; Y) + \mathrm{I}(X; Z|Y)$.
5. $\mathrm{I}(X; Y) = 0$ *iff X, Y are independent.*
6. *If X, Y are jointly independent of Z, then* $\mathrm{I}(X; Y|Z) = \mathrm{I}(X; Y)$.
7. *For any positive integers n and $m \le n/2$,* $\sum_{i=0}^{m} \binom{n}{i} \le 2^{nH_2(m/n)}$.

We will also use the following theorems:

THEOREM 2 (FANO'S INEQUALITY). *Let X be a binary random variable, and let Y be any random variable on a domain $\mathcal{Y}$. Let $f : \mathcal{Y} \rightarrow \{0,1\}$ be a prediction function, which tries to predict the value of X based on an observation of Y. Let $\delta \stackrel{\mathrm{def}}{=} \Pr(f(Y) \neq X)$ be the error probability of the prediction function. Then, $H_2(\delta) \geq \mathrm{H}(X|Y)$.*

THEOREM 3. *Let $C \subseteq \{0,1\}^*$ be a finite prefix-free code (i.e., no codeword in C is a prefix of any other codeword in C). Let X be a random variable corresponding to a uniformly chosen codeword in C. Then, $\mathrm{H}(X) \leq E(|X|)$.*

2.2 Quantum computation

We explain the standard notation of quantum computing and describe the basic notions that will be useful in this paper. For more details we refer the reader to the textbook of Nielsen and Chuang [15].

Let H denote a 2-dimensional complex vector space, equipped with the standard inner product. We pick an orthonormal basis for this space, label the two basis vectors $|0\rangle$ and $|1\rangle$, and for simplicity identify them with the vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively. A *qubit* is a unit length vector in this space, and so can be expressed as a linear combination of the basis states:

$$\alpha_0|0\rangle + \alpha_1|1\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}.$$

Here α_0, α_1 are complex *amplitudes*, and $|\alpha_0|^2 + |\alpha_1|^2 = 1$.

An *m-qubit system* is a unit vector in the m-fold tensor space $H \otimes \cdots \otimes H$. The 2^m basis states of this space are the m-fold tensor products of the states $|0\rangle$ and $|1\rangle$. For example, the basis states of a 2-qubit system are the four 4-dimensional unit vectors $|0\rangle \otimes |0\rangle$, $|0\rangle \otimes |1\rangle$, $|1\rangle \otimes |0\rangle$, and $|1\rangle \otimes |1\rangle$. We abbreviate, e.g., $|1\rangle \otimes |0\rangle$ to $|0\rangle|1\rangle$, or $|1,0\rangle$, or $|10\rangle$, or even $|2\rangle$ (since 2 is 10 in binary). With these basis states, an m-qubit state $|\phi\rangle$ is a 2^m-dimensional complex unit vector

$$|\phi\rangle = \sum_{i \in \{0,1\}^m} \alpha_i |i\rangle.$$

We use $\langle\phi| = |\phi\rangle^*$ to denote the conjugate transpose of the vector $|\phi\rangle$, and $\langle\phi|\psi\rangle = \langle\phi| \cdot |\psi\rangle$ for the inner product between states $|\phi\rangle$ and $|\psi\rangle$. These two states are *orthogonal* if $\langle\phi|\psi\rangle = 0$. The *norm* of $|\phi\rangle$ is $\||\phi\|| = \sqrt{\langle\phi|\phi\rangle}$.

Let $|\phi\rangle$ be an m-qubit state and $B = \{|b_1\rangle, \ldots, |b_{2^m}\rangle\}$ an orthonormal basis of the m-qubit space. A measurement of the state $|\phi\rangle$ in the B basis means that we apply the projection operators $P_i = |b_i\rangle\langle b_i|$ to $|\phi\rangle$. The resulting quantum state is $|b_i\rangle$ with probability $p_i = |\langle\phi|b_i\rangle|^2$.

3. THE HIDDEN MATCHING PROBLEM AND COMPLETE PROBLEMS FOR ONE-WAY COMMUNICATION

Kremer [11] defined a complete problem for quantum one-way communication complexity of Boolean functions. This problem was also considered by Raz [16].

The Problem $\mathcal{P}_0(\theta)$: Alice gets as input a unit vector $\mathbf{x} \in \mathcal{R}^n$. Bob gets as input two orthogonal vector-spaces $M_0, M_1 \subseteq \mathcal{R}^n$ of dimension $n/2$ each. Bob's goal is to output 0 if $\mathbf{x}$ is of distance $\leq \theta$ from M_0 and 1 if $\mathbf{x}$ is of distance $\leq \theta$ from M_1, (and any answer otherwise).

This problem is complete for the class of *Boolean* functions whose quantum one-way communication complexity is *polylog*(n).

We generalize this problem for the case of non-Boolean functions $f : X \times Y \rightarrow [n]$.

The Problem $\mathcal{Q}_0(\theta)$: Alice gets as input a unit vector $\mathbf{x} \in \mathcal{R}^n$. Let $M_0, M_1 \subseteq \mathcal{R}^n$ be two orthogonal vector-spaces of size $n/2$. Bob gets as input an orthonormal basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of $\mathcal{R}^n$ with the property that the basis vectors $\mathbf{b}_i$'s can be partitioned into a basis for M_0 and a basis for M_1. However, this partition is unknown to Alice and Bob. Bob's goal is to output a value in $\{i \mid \mathbf{b}_i \in M_0\}$ if $\mathbf{x}$ is of distance $\leq \theta$ from M_0 and a value in $\{i \mid \mathbf{b}_i \in M_1\}$ if $\mathbf{x}$ is of distance $\leq \theta$ from M_1, (and any answer otherwise).

We can prove the following:

THEOREM 4. *For any $0 \leq \theta < 1/\sqrt{2}$, the problem $\mathcal{Q}_0(\theta)$ is complete for the class of relations $\mathcal{R} \subseteq X \times Y \times Z$ whose quantum one-way communication complexity is polylog(n).*

PROOF. We first show that the problem can be solved efficiently by a quantum one-way Protocol. Alice just encodes the unit vector $\mathbf{x} \in \mathbf{R}^n$ by $\log n$ qubits and sends this to Bob. Bob measures in the basis B. If $\mathbf{x}$ is of distance $\leq \theta$ from M_0 the answer will be in $\{i|\mathbf{b}_i \in M_0\}$ with probability $\geq 1 - \theta^2$. If $\mathbf{x}$ is of distance $\leq \theta$ from M_1 the answer will be an $\{i|\mathbf{b}_i \in M_1\}$ with probability $\geq 1 - \theta^2$.

Next, we want to reduce any problem $\mathcal{R} \subseteq X \times Y \times Z$ with one-way communication d to the problem $\mathcal{Q}_0(\theta)$ with input size $n = 2^{O(d)}$ and $0 < \theta < 1/\sqrt{2}$. In any protocol Alice applies some unitary operation U on the initial state $|0\rangle$ that depends on the input $\mathbf{x}$ and Bob measures in some basis B of $\mathbf{R}^n$. Let $\mathbf{x}'$ be the unit vector $U|0\rangle$. Since the communication is d, $\mathbf{x}'$ lies in $\mathbf{R}^{2^{O(d)}}$ (see also page 9, first comment in [16]). Then, $\mathbf{x}'$ and the basis B can be used as the inputs in $\mathcal{Q}_0(\theta)$. $\square$

Hence, the lower bound we obtain for the problem HM_n translates into a lower bound for the complete problem $\mathcal{Q}_0(\theta)$. Also, the lower bounds we prove for the Boolean Hidden Matching Problem translate into bounds for the complete Boolean problem $\mathcal{P}_0(\theta)$.

4. THE QUANTUM UPPER BOUND

We present a quantum protocol for the hidden matching problem with communication complexity of $\log n$ qubits. Let $\mathbf{x} = x_1 \ldots x_n$ be Alice's input and $M \in \mathcal{M}_n$ be Bob's input.

QUANTUM PROTOCOL FOR HM_n

1. Alice sends the state $|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} |i\rangle$.

2. Bob performs a measurement on the state $|\psi\rangle$ in the orthonormal basis
$B = \{\frac{1}{\sqrt{2}}(|k\rangle \pm |\ell\rangle) \mid (k, \ell) \in M\}$.

The probability that the outcome of the measurement is a basis state $\frac{1}{\sqrt{2}}(|k\rangle + |\ell\rangle)$ is

$$|\langle\psi|\frac{1}{\sqrt{2}}(|k\rangle + |\ell\rangle)\rangle|^2 = \frac{1}{2n}((-1)^{x_k} + (-1)^{x_\ell})^2.$$

This equals to $2/n$ if $x_k \oplus x_\ell = 0$ and 0 otherwise. Similarly for the states $\frac{1}{\sqrt{2}}(|k\rangle - |\ell\rangle)$ we have that $|\langle\psi|\frac{1}{\sqrt{2}}(|k\rangle - |\ell\rangle)\rangle|^2$ is 0 if $x_k \oplus x_\ell = 0$ and $2/n$ if $x_k \oplus x_\ell = 1$. Hence, if the outcome of the measurement is a state $\frac{1}{\sqrt{2}}(|k\rangle + |\ell\rangle)$ then Bob knows with certainty that $x_k \oplus x_\ell = 0$ and outputs $\langle k, \ell, 0\rangle$. If the outcome is a state $\frac{1}{\sqrt{2}}(|k\rangle - |\ell\rangle)$ then Bob knows with certainty that $x_k \oplus x_\ell = 1$ and hence outputs $\langle k, \ell, 1\rangle$. Note that the measurement depends only on Bob's input and that the algorithm is 0-error.

Remark : As mentioned above, the inspiration for this problem comes from Locally Decodable Codes. A 2-query Locally Decodable Code is a code $C : \{0,1\}^n \to \{0,1\}^m$ with the property that for every index $k \in [n]$ there exists a matching M_k on the coordinates of $C(\mathbf{x})$, such that for every pair $(i,j) \in M_k$, $x_k = f_C(k, i, j, C(\mathbf{x})_i \oplus C(\mathbf{x})_j)$; here, f_C is a fixed function depending only on the code C. We can cast this problem as a communication problem, by letting Alice have the codeword $C(\mathbf{x})$ and Bob have the index k and the corresponding matching M_k, and the goal is to output $x_k = f_C(k, i, j, C(\mathbf{x})_i \oplus C(\mathbf{x})_j)$ for some $(i,j) \in M_k$. This gives rise to our Hidden Matching Problem. The above mentioned quantum algorithm is based on the fact that a uniform superposition of $C(\mathbf{x})$ is sufficient to compute the parity of some pair in the matching. The same property was used by Kerenidis and de Wolf [9] to prove a lower bound on the length of classical 2-query Locally Decodable Codes.

5. THE RANDOMIZED LOWER BOUND

THEOREM 5. *Any one-way randomized protocol for computing* HM_n *with error probability less than* $1/16$ *requires* $\Omega(\sqrt{n})$ *bits of communication.*

PROOF. Using Yao's Lemma [20], in order to prove the lower bound, it suffices to construct a "hard" distribution μ over instances of HM_n, and prove a distributional lower bound w.r.t. deterministic one-way protocols. We define μ as follows: let $\mathbf{X}$ be a uniformly chosen bitstring in $\{0,1\}^n$; let $\mathbf{M}$ be an independent and uniformly chosen perfect matching in $\mathcal{M}$, where $\mathcal{M}$ is any set of $m = \Omega(n)$ pairwise edge-disjoint matchings. (For example, let $m = n/2$ and $\mathcal{M} = \{M_1, M_2, \ldots, M_m\}$, where M_i is defined by the edge set $\{(j, m + (j+i) \bmod m) \mid j = 0, \ldots, m-1\}$.)

In the proof we use the following simple fact, which is proved using Markov's inequality:

PROPOSITION 1. *Let X be a random variable with values in the interval* $[0,1]$. *If $E[X] \geq 1 - \epsilon$, then for any $t \geq 1$,* $\Pr[X \geq 1 - t\epsilon] \geq 1 - 1/t$.

Let Π be a deterministic protocol for this problem with (distributional) error $\delta < 1/16$ with respect to μ.

Define the $2^n \times m$ protocol matrix P whose rows and columns are indexed by the inputs $\mathbf{x}$ to Alice and inputs M to Bob, respectively. The entry at $(\mathbf{x}, M)$ is the output $\langle i, j, b\rangle$ of Π on $(\mathbf{x}, M)$. We will assume, without loss of generality, that $(i,j) \in M$, thus an error occurs at $(\mathbf{x}, M)$ only when $b \neq x_i \oplus x_j$. When this happens we say that the entry of P at $(\mathbf{x}, M)$ is incorrect.

For any possible message τ of Alice, let S_τ denote the set of Alice's inputs on which she sends τ. The *weight* of τ is defined to be the fraction $|S_\tau|/2^n$. We call τ *good*, if at least $1 - 2\delta$ fraction of the entries in the submatrix $S_\tau \times \mathcal{M}$ of P are correct. By Proposition 1, the total weight of good τ's is at least $1/2$. For any such τ, we will show that its weight is at most $2^{-\Omega(\sqrt{n})}$. It would follow that the number of good τ's is at least $(1/2)/2^{-\Omega(\sqrt{n})} = 2^{\Omega(\sqrt{n})}$, and therefore the communication cost of Π has to be at least $\Omega(\sqrt{n})$. For the rest of the proof fix such a τ.

Each entry in P consists of an edge and a bit.t Therefore, any row of P defines a graph obtained by taking the m edges in that row, and a vector in $\{0,1\}^m$ corresponding to the bits in that row. Since Π is a one-round protocol, the output depends only on τ and Bob's input. It follows that the rows of P corresponding to inputs in S_τ are associated with the same graph $G = G_\tau$ and the same vector $\mathbf{u} = \mathbf{u}_\tau$.

Recall that since τ is good, the fraction of correct entries in the submatrix $S_\tau \times \mathcal{M}$ is at least $1 - 2\delta$. By another application of Proposition 1, it follows that for at least half of the columns in this submatrix, the fraction of correct entries in any such column is at least $1 - 4\delta$. Let G' denote the set of edges associated with these columns. Thus, $|G'| \geq m/2$.

We next show that G' contains a forest with $\Omega(\sqrt{m}) = \Omega(\sqrt{n})$ edges. Let $C_1, C_2, \ldots, C_s$ be the connected components of G', and let $b_1, b_2, ..., b_s$ be the number of edges they have $(b_1 + b_2 + ... + b_s = |G'|)$. C_i has at least $\Omega(\sqrt{b_i})$ nodes, and thus has a spanning tree with $\Omega(\sqrt{b_i})$ edges. Therefore, G' contains a forest F with at least $\sum_i \Omega(\sqrt{b_i}) = \Omega(\sqrt{m})$ edges, using the fact that $\sqrt{u} + \sqrt{v} \geq \sqrt{u+v}$.

Consider now the submatrix $S_\tau \times F$. Since F is a subgraph of G', at least $1 - 4\delta$ of the entries in this submatrix are correct. Thus, by Proposition 1, at least $1/2$ of the rows have the property that at least $1 - 8\delta$ fraction of the entries in each such row are correct. Call these rows S'_τ.

Let N denote the $n \times |F|$ vertex-edge incidence matrix of F, and let $\mathbf{v} \in \{0,1\}^{|F|}$ denote the projection of $\mathbf{u}$ on F. For any $\mathbf{x} \in \{0,1\}^n$, the vector $\mathbf{x}N$ taken over $GF[2]$ is of length $|F|$. If the i-th coordinate corresponds to the edge (j,k), then the value of $\mathbf{x}N$ at this coordinate equals $x_j \oplus x_k$. Any input $\mathbf{x} \in S'_\tau$ therefore satisfies $h(\mathbf{x}N, \mathbf{v}) \leq 8\delta$, where $h(\cdot,\cdot)$ denotes the relative Hamming distance. If $W = \{\mathbf{w} : h(\mathbf{w}, \mathbf{v}) \leq 8\delta\}$, it follows that $S'_\tau \subseteq \bigcup_{\mathbf{w} \in W} \{\mathbf{x} : \mathbf{x}N = \mathbf{w}\}$.

Since F is a forest, the columns of N are linearly independent over $GF[2]$, implying that the null space $\{\mathbf{z} : \mathbf{z}N = 0\}$ has dimension $n - |F|$. Therefore, for any $\mathbf{w} \in W$, the number of solutions $\mathbf{x}$ such that $\mathbf{x}N = \mathbf{w}$ is at most $2^{n-|F|}$. By the estimate given in Theorem 1, part (8), $|W| \leq 2^{|F| \cdot H_2(8\delta)}$, where $H_2(\cdot)$ denotes the binary entropy function. Thus, $|S'_\tau| \leq 2^{n - |F|(1 - H_2(8\delta))}$.

Recall that $|F| = \Omega(\sqrt{n})$ and that $|S'_\tau| \geq |S_\tau|/2$. We conclude that the weight of S_τ, $|S_\tau|/2^n$, is at most $2|S'_\tau|/2^n \leq 2^{1-|F|(1-H_2(8\delta))} = 2^{-\Omega(\sqrt{n})}$, as needed. □

An $O(\sqrt{n})$ public-coin randomized protocol for HM_n is as follows. Alice uses the public coins to pick $\sqrt{n}$ locations in $[n]$ and sends the corresponding bits to Bob. Standard calculations show that these bits include the end-points of at least one matching edge with constant probability. Thus:

THEOREM 6. *The randomized one-way communication complexity of* HM_n *is* $\Theta(\sqrt{n})$.

6. AN EXPONENTIAL SEPARATION FOR SIMULTANEOUS MESSAGES

Recall that in the model of Simultaneous Messages (SM), Alice and Bob both send a single message to a Referee, after which he computes the output. We prove an exponential separation in this model between quantum and public-coin randomized communication complexity. To this end, we use a restricted version of the Hidden Matching problem, in which Bob's input is not any perfect matching on n vertices, but rather only one of $m = \Omega(n)$ fixed matchings (Alice, Bob, and referee know this collection of m matchings a priori). We denote this problem by $\mathrm{HM}_n^{\mathrm{small}}$.

The lower bound we proved for HM_n in the model of one-way communication (Theorem 5) is in fact a lower bound for $\mathrm{HM}_n^{\mathrm{small}}$. This lower bound holds also in the SM model since this model is no more powerful than one-way communication. On the other hand, the problem is still easy in the quantum case. Bob sends the index of his matching to the Referee using $\log n$ bits and Alice sends a superposition of her input string using $\log n$ qubits, similarly to the one-way protocol. Since the referee knows Bob's matching, he can perform the same measurement Bob performed in the one-way protocol and compute the XOR of some pair in the matching.

7. THE COMPLEXITY OF BOOLEAN HIDDEN MATCHING

The $O(\log n)$ qubit quantum protocol for HM_n can be tweaked to solve also BHM_n: after obtaining the value $\langle k, \ell, c\rangle$ from that protocol, where (k, ℓ) is the i-th pair in Bob's input matching M, Bob outputs $w_i \oplus c$. Note that if $c = x_k \oplus x_\ell$, then $w_i \oplus c$ equals the desired bit b.

7.1 Lower bound for 0-error protocols

In order to prove the lower bound for 0-error protocols, we note the following characterization of 0-error randomized one-way communication complexity of partial functions. Let $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1, *\}$ be a partial Boolean function. We say that the input (x, y) is *legal*, if $f(x, y) \neq *$. A protocol for f is required to be correct only on legal inputs; it is allowed to output arbitrary answers on illegal inputs. The *confusion graph* G_f of f is a graph whose vertex set is $\mathcal{X}$; (x, x') is an edge in G_f if and only if there exists a y such that both (x, y) and (x', y) are legal inputs and $f(x, y) \neq f(x', y)$.

It is known [12] that the deterministic one-way communication complexity of f is $\log \chi(G_f) + O(1)$, where $\chi(G_f)$ is the chromatic number of the graph G_f. We will obtain a lower bound on the 0-error randomized one-way communication complexity via another measure on G_f. For any graph $G = (V, E)$, let

$$\theta(G) = \max_{W \subseteq V} \frac{|W|}{\alpha(G_W)},$$

where G_W is the subgraph of G induced on W and $\alpha(G_W)$ is the independence number of G_W. It is easy to see that $\chi(G) \geq \theta(G)$. The following theorem shows that $\theta(G_f)$ is a lower bound on the 0-error communication complexity of f.

THEOREM 7. *The 0-error randomized one-way communication complexity of any partial Boolean function f is at least $\log \theta(G_f)$.*

PROOF. Let $G_f = (V, E)$ and let $W \subseteq V$ achieve the maximum for $\theta(G_f)$. Define μ to be the uniform distribution on W.

Suppose Π is a randomized 0-error one-way protocol for f with public randomness R, and whose cost is $c + 1$ (Bob just outputs a bit which is the last bit of the transcript). Let $A(\mathbf{x}, R)$ be the message sent by Alice on input $\mathbf{x}$, and let $B(\tau, y, R)$ be the output of the protocol given by Bob on input y when the message sent by Alice is τ. For any legal input (x, y), we have $E[|A(x, R)|] \leq c$, and $\Pr[B(A(x, R), y, R) = f(x, y)] = 1$.

Let X be a random input for Alice whose distribution is μ. Then $E[|A(X, R)|] \leq c$, where the randomness is now over both X and R. Therefore, there exists a choice r^* for R such that $E[|A(X, r^*)|] \leq c$. Define a deterministic protocol where $A'(x) = A(x, r^*)$ and $B'(\tau, y) = B(\tau, y, r^*)$. Note that this protocol correctly computes f and $E[|A'(X)|] \leq c$. Let T be the set of messages sent by Alice in this new protocol. For any message $\tau \in T$, define $S_\tau = \{x \in W : A'(x) = \tau\}$. By the definition of G_f, it follows that S_τ is an independent set, so $|S_\tau| \leq \alpha(G_W)$. Therefore, the entropy of the random variable $A'(X)$ satisfies:

$$\begin{aligned} \mathrm{H}(A'(X)) &= \sum_{\tau \in T} \frac{|S_\tau|}{|W|} \log\Big(\frac{|W|}{|S_\tau|}\Big) \qquad (1) \\ &\geq \sum_{\tau \in T} \frac{|S_\tau|}{|W|} \log\Big(\frac{|W|}{\alpha(G_W)}\Big) = \log \theta(G_f), \end{aligned}$$

because the S_τ's partition W.

Finally, if we assume that the messages are prefix-free (which can be achieved with a constant factor blow-up in the communication cost), then $E[|A'(X)|] \geq \mathrm{H}(A'(X))$ (Theorem 3). It follows from Equation 1 that $c \geq \log \theta(G_f)$. $\square$

We use this characterization to prove the lower bound for BHM_n:

THEOREM 8. *Let $n = 4p$, where p is prime. Then, the 0-error randomized one-way communication complexity of BHM_n is $\Omega(n)$.*

PROOF. Let f denote the partial function BHM_n. The vertex set of the confusion graph G_f is $\{0,1\}^n$. We next show that $(\mathbf{x}, \mathbf{x}')$ is an edge in G_f if and only if the Hamming distance between $\mathbf{x}$ and $\mathbf{x}'$ is exactly $n/2$.

Suppose $(\mathbf{x}, \mathbf{x}')$ is an edge in G_f. Therefore, there exists a matching M and a vector $\mathbf{w}$, so that $M\mathbf{x} \oplus \mathbf{w} = \mathbf{0}$ and $M\mathbf{x}' \oplus \mathbf{w} = \mathbf{1}$, or vice versa. That means that for every edge $(i, j) \in M$, $x_i \oplus x_j \neq x'_i \oplus x'_j$, and thus $\mathbf{x}, \mathbf{x}'$ agree on one of the position i, j and disagree on the other. Hence, the Hamming distance between $\mathbf{x}$ and $\mathbf{x}'$ is exactly $n/2$. Conversely, given two strings $\mathbf{x}, \mathbf{x}'$ of Hamming distance $n/2$, let M be any matching between the positions on which $\mathbf{x}, \mathbf{x}'$ agree and the positions on which they disagree. Let $\mathbf{w} = M\mathbf{x}$. Clearly, $M\mathbf{x} \oplus \mathbf{w} = \mathbf{0}$. For each edge (i, j) in M we have $x_i \oplus x_j \neq x'_i \oplus x'_j$, and therefore $M\mathbf{x}' \oplus \mathbf{w} = \mathbf{1}$, implying $(\mathbf{x}, \mathbf{x}')$ is an edge in G_f.

If $n/2$ is odd, G_f is the bipartite graph between the even and odd parity vertices. Therefore, G_f is 2-colorable, implying that f has a $O(1)$ protocol (Alice just sends the parity of her input). We will show that the situation changes dramatically when n is a multiple of 4.

PROPOSITION 2 (FRANKL AND WILSON [7]). *Let $m = 4p - 1$, where p is prime. Define the graph $G = (V, E)$ where*

$V = \{A \subseteq [m] : |A| = 2p-1\}$, *and* $(A,B) \in E$ *if and only if* $|A \cap B| = p-1$. *Then,*

$$\alpha(G) \le \sum_{i=0}^{p-1} \binom{m}{i}.$$

Let $m = 4p-1$ and let G be the graph defined by Proposition 2. We claim that G is isomorphic to a vertex-induced subgraph of the confusion graph G_f: for every vertex A in G, the corresponding vertex in G_f is the characteristic vector of the set $A \cup \{4p\}$. Let V denote the vertex set of G; it follows that $\theta(G_f) \ge |V|/\alpha(G)$.

We have, $|V| = \binom{m}{2p-1} \approx 2^m/\sqrt{m}$, and by Proposition 2, $\alpha(G) \le 2^{mH_2(\gamma)}$, where H_2 is the binary entropy function and $\gamma = (p-1)/(4p-1) \le 1/4$. The result now follows from Theorem 7. $\square$

7.2 Lower bound for linear randomized protocols

THEOREM 9. *Let* n *be a positive integer multiple of 4, and let* $0 < \delta < 1/2$ *be a constant bounded away from* $1/2$. *Then, any* δ*-error public-coin one-way linear protocol for* BHM_n *requires* $\Omega(\sqrt[3]{n \log n})$ *bits of communication.*

PROOF. Using Yao's Lemma [20], in order to prove the lower bound, it suffices to construct a "hard" distribution μ over instances of BHM_n, and prove a distributional lower bound w.r.t. deterministic one-way linear protocols. We define μ as follows: let $\mathbf{X}$ be a uniformly chosen bitstring in $\{0,1\}^n$; let $\mathbf{M}$ be a uniformly chosen perfect matching in $\mathcal{M}_n$; and let B be a uniformly chosen bit. $\mathbf{W}$ is a random bitstring in $\{0,1\}^{n/2}$, defined as $\mathbf{W} \stackrel{\text{def}}{=} \mathbf{MX} \oplus \mathbf{B}$ (recall that $\mathbf{B}$ is the vector all of whose entries are B).

Let Π be any deterministic one-way linear protocol that has error probability of at most δ when solving BHM_n on inputs drawn according to μ. Let c be the communication cost of Π.

Since Π is deterministic, one-way, and linear, there exists a fixed $c \times n$ Boolean matrix A, such that the message of Π on any input $\mathbf{x}$ is $A\mathbf{x}$. By adding at most one bit to the communication cost of Π, we can assume $\mathbf{1}$ is one of the rows of A. We also assume, without loss of generality, that A has a full row rank, because otherwise Alice sends redundant information, which Bob can figure out by himself.

We assume c satisfies $c^3/\log c \le 3n/4$, since, otherwise, $c \ge \Omega(\sqrt[3]{n \log n})$, and we are done.

For a matrix T, we denote by $\mathrm{sp}(T)$ the span of the row vectors of T over the field $GF(2)$. Clearly, for any matrix T, $\mathbf{0} \in \mathrm{sp}(T)$. In particular, $\mathbf{0} \in \mathrm{sp}(M) \cap \mathrm{sp}(A)$, for any matching $M \in \mathcal{M}_n$ (recall that we view a matching M as an $\frac{n}{2} \times n$ edge-vertex incidence matrix). By our assumption about A, $\mathbf{1} \in \mathrm{sp}(A)$. Since M is a perfect matching, the sum of its rows is $\mathbf{1}$, thus $\mathbf{1} \in \mathrm{sp}(M)$. We conclude that for any M, $\{\mathbf{0}, \mathbf{1}\} \subseteq \mathrm{sp}(M) \cap \mathrm{sp}(A)$. Let Z be an indicator random variable of the event $\{\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}(A) = \{\mathbf{0}, \mathbf{1}\}\}$, meaning that $\mathbf{0}$ and $\mathbf{1}$ are the only vectors in the intersection of the spans.

In the protocol Π, Bob observes values of the random variables $A\mathbf{X}, \mathbf{M}$, and $\mathbf{W}$ and uses them to predict the random variable B with error probability δ. Therefore, by Fano's inequality (Theorem 2),

$$H_2(\delta) \;\ge\; \mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}). \tag{2}$$

Since conditioning reduces entropy,

$$\begin{aligned} \mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}) \;&\ge\; \mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}, Z) \\ &= \mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}, Z=1) \cdot \Pr(Z=1) \\ &\quad + \mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}, Z=0) \cdot \Pr(Z=0) \\ &\ge \mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}, Z=1) \cdot \Pr(Z=1). \end{aligned} \tag{3}$$

The following two lemmas bound the two factors in the last expression:

LEMMA 1. $\mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}, Z=1) = 1$.

LEMMA 2. $\Pr(Z=1) \ge 1 - O(\frac{c^3}{n \log c})$.

The proofs of the Lemma 1 and 2 are provided below. Let us first show how the two lemmas derive the theorem. By combining Equations 2 and 3, and Lemmas 1 and 2, we have:

$$H_2(\delta) \;\ge\; 1 - O(\frac{c^3}{n \log c}).$$

Therefore,

$$\begin{aligned} c \;&\ge\; \Omega(\sqrt[3]{n(1 - H_2(\delta)) \cdot \log(n(1 - H_2(\delta)))}) \\ &= \Omega(\sqrt[3]{n \log n}), \end{aligned}$$

since $H_2(\delta)$ is a constant bounded away from 1. This completes the proof of the theorem. $\square$

PROOF OF LEMMA 1. Recall that we assume $\mathbf{1}$ is one of the rows of A and that A has a full row rank. Let A' be the submatrix of A consisting of all the rows of A, except $\mathbf{1}$. Clearly, $\mathrm{sp}(A') \subseteq \mathrm{sp}(A)$ and $\mathbf{1} \notin \mathrm{sp}(A')$. It follows that the event $\{\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}(A) = \{\mathbf{0}, \mathbf{1}\}\}$ is the same as the event $\{\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}(A') = \{\mathbf{0}\}\}$. Thus, from now on we will think of Z as an indicator random variable of the latter.

Observe that since n is a multiple of 4, the parity of the bits of $\mathbf{w}$ always equals to the parity of the bits of $\mathbf{x}$. It follows that in the pair of random variables $(A\mathbf{X}, \mathbf{W})$ the same information (that is, the random variable $\mathbf{1} \cdot \mathbf{X}$) is repeated twice. We can therefore rewrite $\mathrm{H}(B \mid A\mathbf{X}, \mathbf{M}, \mathbf{W}, Z=1)$ as $\mathrm{H}(B \mid A'\mathbf{X}, \mathbf{M}, \mathbf{W}, Z=1)$.

By the definition of mutual information,

$$\begin{aligned} &\mathrm{H}(B \mid A'\mathbf{X}, \mathbf{M}, \mathbf{W}, Z=1) \\ &\quad = \mathrm{H}(B \mid \mathbf{M}, \mathbf{W}, Z=1) - \mathrm{I}(B \,;\, A'\mathbf{X} \mid \mathbf{M}, \mathbf{W}, Z=1). \end{aligned}$$

The next proposition shows that the random variables $B, \mathbf{M}$, and $\mathbf{W}$ are mutually independent given the event $\{Z=1\}$, implying that $\mathrm{H}(B \mid \mathbf{M}, \mathbf{W}, Z=1) = \mathrm{H}(B|Z=1) = \mathrm{H}(B) = 1$. Thus, in order to prove the lemma it would suffice to show that $\mathrm{I}(B \,;\, A'\mathbf{X} \mid \mathbf{M}, \mathbf{W}, Z=1) = 0$.

PROPOSITION 3. *The random variables* $B, \mathbf{M}$, *and* $\mathbf{W}$ *are mutually independent, given the event* $\{Z=1\}$.

PROOF. We will show the random variables $B, \mathbf{M}$, and $\mathbf{W}$ are mutually independent unconditionally. This independence would then hold even given the event $\{Z=1\}$, because this event is a function of $\mathbf{M}$ only.

The random variables B and $\mathbf{M}$ are independent, by definition. Let M be any value of the random variable $\mathbf{M}$, and let b be any value of the random variable B. In order to show the desired independence, we need to prove that for any possible value $\mathbf{w}$ of $\mathbf{W}$, $\Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b) = \Pr(\mathbf{W} = \mathbf{w})$.

Using conditional probability, we can rewrite $\Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b)$ as follows:

$$\Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b) \;=\; \sum_{\mathbf{x} \in \{0,1\}^n} \Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b, \mathbf{X} = \mathbf{x}) \cdot \Pr(\mathbf{X} = \mathbf{x} \mid \mathbf{M} = M, B = b).$$

Since $\mathbf{X}, \mathbf{M}$,and B are mutually independent by definition, then $\Pr(\mathbf{X} = \mathbf{x} \mid \mathbf{M} = M, B = b) = \Pr(\mathbf{X} = \mathbf{x}) = 1/2^n$. $\Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b, \mathbf{X} = \mathbf{x}) = 1$ only if $\mathbf{w} = M\mathbf{x} \oplus \mathbf{b}$, and it is 0 otherwise. The number of $\mathbf{x}$'s that satisfy this condition is the number of solutions to the linear system $M\mathbf{x} = \mathbf{w} \oplus \mathbf{b}$ over Z_2^n. Since M is an $\frac{n}{2} \times n$ matrix that has a full row rank, this number is $2^{n/2}$. Therefore, $\Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b) = 2^{n/2}/2^n = 1/2^{n/2}$.

Consider now the quantity $\Pr(\mathbf{W} = \mathbf{w})$. Using conditional probability we can rewrite it as:

$$\Pr(\mathbf{W} = \mathbf{w}) \;=\; \sum_{M,b} \Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b) \cdot \Pr(\mathbf{M} = M, B = b).$$

We already proved that for all M and b, $\Pr(\mathbf{W} = \mathbf{w} \mid \mathbf{M} = M, B = b) = 1/2^{n/2}$. Therefore, also $\Pr(\mathbf{W} = \mathbf{w}) = 1/2^{n/2}$, completing the proof. $\square$

Next we prove $\mathrm{I}(B \,;\, A'\mathbf{X} \mid \mathbf{M}, \mathbf{W}, Z = 1) = 0$. By the chain rule for mutual information,

$$\begin{aligned}&\mathrm{I}(B, \mathbf{M}, \mathbf{W} \,;\, A'\mathbf{X} \mid Z = 1)\\ &= \mathrm{I}(\mathbf{M}, \mathbf{W} \,;\, A'\mathbf{X} \mid Z = 1) + \mathrm{I}(B \,;\, A'\mathbf{X} \mid \mathbf{M}, \mathbf{W}, Z = 1).\end{aligned}$$

Since mutual information is always a non-negative quantity, it would thus suffice to show that $\mathrm{I}(B, \mathbf{M}, \mathbf{W} \,;\, A'\mathbf{X} \mid Z = 1) = 0$.

The function $f(b, M, \mathbf{w}) = (b, M, \mathbf{w} \oplus \mathbf{b})$ is a 1-1 function. Note that $f(B, \mathbf{M}, \mathbf{W}) = (B, \mathbf{M}, \mathbf{W} \oplus \mathbf{B}) = (B, \mathbf{M}, \mathbf{MX})$. Therefore, by the data processing inequality (applied using both f and f^{-1}), we have:

$$\mathrm{I}(B, \mathbf{M}, \mathbf{W} \,;\, A'\mathbf{X} \mid Z = 1) \;=\; \mathrm{I}(B, \mathbf{M}, \mathbf{MX} \,;\, A'\mathbf{X} \mid Z = 1).$$

Using again the chain rule for mutual information we have:

$$\begin{aligned}&\mathrm{I}(B, \mathbf{M}, \mathbf{MX} \,;\, A'\mathbf{X} \mid Z = 1) \;=\; \\ &\mathrm{I}(B, \mathbf{M} \,;\, A'\mathbf{X} \mid Z = 1) + \mathrm{I}(\mathbf{MX} \,;\, A'\mathbf{X} \mid B, \mathbf{M}, Z = 1).\end{aligned} \tag{4}$$

We next show that each of the above mutual information quantities is 0. By the definition of the input distribution μ, the random variables $B, \mathbf{M}$, and $\mathbf{X}$ are mutually independent. This holds even given the event $\{Z = 1\}$, because the latter is a function of $\mathbf{M}$ only. It follows that also $B, \mathbf{M}$, $A'\mathbf{X}$ are mutually independent given the event $\{Z = 1\}$, and thus $\mathrm{I}(B, \mathbf{M} \,;\, A'\mathbf{X} \mid Z = 1) = 0$.

As for the second mutual information quantity on the RHS of Equation 4, we use again the independence of $B, \mathbf{M}$, and $A'\mathbf{X}$ given $\{Z = 1\}$ to derive $\mathrm{I}(\mathbf{MX} \,;\, A'\mathbf{X} \mid B, \mathbf{M}, Z = 1) = \mathrm{I}(\mathbf{MX} \,;\, A'\mathbf{X} \mid \mathbf{M}, Z = 1)$. The following proposition proves that for any matching M satisfying the condition indicated by the event $\{Z = 1\}$, the random variables $M\mathbf{X}$ and $A'\mathbf{X}$ are independent. It then follows that $\mathrm{I}(\mathbf{MX} \,;\, A'\mathbf{X} \mid \mathbf{M}, Z = 1) = 0$.

Proposition 4. *For any matching $M \in \mathcal{M}_n$ satisfying the condition $\mathrm{sp}(M) \cap \mathrm{sp}(A') = \{\mathbf{0}\}$, the random variables $M\mathbf{X}$ and $A'\mathbf{X}$ are independent.*

Proof. Let $\mathbf{z}$ be any possible value for the random variable $M\mathbf{X}$ and let $\mathbf{y}$ be any possible value for the random variable $A'\mathbf{X}$. In order to prove the independence, we need to show that $\Pr(M\mathbf{X} = \mathbf{z} \mid A'\mathbf{X} = \mathbf{y}) = \Pr(M\mathbf{X} = \mathbf{z})$.

M is an $\frac{n}{2} \times n$ Boolean matrix that has a full row rank. Therefore, the number of solutions to the linear system $M\mathbf{x} = \mathbf{z}$ over Z_2^n is exactly $2^{n/2}$. Recall that $\mathbf{X}$ was chosen uniformly at random from Z_2^n. Therefore, $\Pr(M\mathbf{X} = \mathbf{z}) = 1/2^{n/2}$.

By the definition of conditional probability, $\Pr(M\mathbf{X} = \mathbf{z} \mid A'\mathbf{X} = \mathbf{y}) = \Pr(M\mathbf{X} = \mathbf{z}, A'\mathbf{X} = \mathbf{y}) / \Pr(A'\mathbf{X} = \mathbf{y})$. Since A' is a $(c-1) \times n$ Boolean matrix and has a full row rank, the same argument as above shows that $\Pr(A'\mathbf{X} = \mathbf{y}) = 1/2^{n-c+1}$. Let D be an $(\frac{n}{2} + c - 1) \times n$ matrix, which is composed by putting M on top of A'. Since $\mathrm{sp}(M) \cap \mathrm{sp}(A') = \{\mathbf{0}\}$, D is has a full row rank. We thus obtain $\Pr(M\mathbf{X} = \mathbf{z}, A'\mathbf{X} = \mathbf{y}) = \Pr(D\mathbf{X} = (\mathbf{z}, \mathbf{y})) = 1/2^{n/2-c+1}$. Hence, $\Pr(M\mathbf{X} = \mathbf{z} \mid A'\mathbf{X} = \mathbf{y}) = 2^{n/2-c+1}/2^{n-c+1} = 1/2^{n/2} = \Pr(M\mathbf{X} = \mathbf{z})$. The proposition follows. $\square$

This completes the proof of Lemma 1.

Proof of Lemma 2. Denote the event $\{\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}(A) \neq \{\mathbf{0}, \mathbf{1}\}\}$ by E. We would like to prove $\Pr(E) \leq O(c^3/(n \log c))$. For $0 \leq k \leq n$, define $\mathrm{sp}_k(A)$ to be the vectors in $\mathrm{sp}(A)$ whose Hamming weight is k. Define E_k to be the event $\{\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}_k(A) \neq \emptyset\}$. Since $\mathrm{sp}_0(A) = \{\mathbf{0}\}$ and $\mathrm{sp}_n(A) = \{\mathbf{1}\}$, the event E can be rewritten as $\bigvee_{k=1}^{n-1} E_k$. Thus, using the union bound, we can bound the probability of E as follows:

$$\Pr(E) \;\leq\; \sum_{k=1}^{n-1} \Pr(\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}_k(A) \neq \emptyset). \tag{5}$$

Let M be any matching in $\mathcal{M}_n$. Any vector $\mathbf{v}$ in $\mathrm{sp}(M)$, when viewed as a set $S_{\mathbf{v}}$ (i.e., $\mathbf{v}$ is the characteristic vector of $S_{\mathbf{v}}$), is a disjoint union of edges from M. We thus immediately conclude that $\mathbf{v}$ has to have an even Hamming weight. This implies that for all odd $1 \leq k \leq n-1$,

$$\Pr(\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}_k(A) \neq \emptyset) = 0. \tag{6}$$

Consider then an even k, and let $\mathbf{v}$ be any vector in $\mathrm{sp}_k(A)$. If $\mathbf{v}$ belongs to $\mathrm{sp}(M)$, then M can be partitioned into two perfect "sub-matchings": a perfect matching on $S_{\mathbf{v}}$ and perfect matching on $[n] \setminus S_{\mathbf{v}}$. We conclude that the number of matchings M in $\mathcal{M}_n$, for which $\mathbf{v} \in \mathrm{sp}(M)$, is exactly $m_k \cdot m_{n-k}$, where m_ℓ is the number of perfect matchings on ℓ nodes. Note that $m_\ell = \frac{\ell!}{(\ell/2)! 2^{\ell/2}}$, and thus,

$$\Pr(\mathbf{v} \in \mathrm{sp}(\mathbf{M})) \;=\; \frac{m_k \cdot m_{n-k}}{m_n} \;=\; \frac{\binom{\frac{n}{2}}{\frac{k}{2}}}{\binom{n}{k}}.$$

It follows, by the union bound, that for any even k,

$$\Pr(\mathrm{sp}(\mathbf{M}) \cap \mathrm{sp}_k(A) \neq \emptyset) \;\leq\; |\mathrm{sp}_k(A)| \cdot \frac{\binom{\frac{n}{2}}{\frac{k}{2}}}{\binom{n}{k}}. \tag{7}$$

Since $\mathbf{1} \in \mathrm{sp}(A)$, then $|\mathrm{sp}_k(A)| = |\mathrm{sp}_{n-k}(A)|$, for all $0 \leq k \leq n$. Combining this and Equations 5, 6, and 7, it would thus suffice to prove the following:

$$\sum_{j=1}^{n/4} |\mathrm{sp}_{2j}(A)| \cdot \frac{\binom{\frac{n}{2}}{j}}{\binom{n}{2j}} \;\leq\; O(\frac{c^3}{n \log c}). \tag{8}$$

We start by bounding the ratio in each of the terms:

$$
\begin{aligned}
\frac{\binom{\frac{n}{2}}{j}}{\binom{n}{2j}} &= \frac{(\frac{n}{2})! \cdot (2j)! \cdot (n-2j)!}{(\frac{n}{2}-j)! \cdot j! \cdot n!} \\
&= \frac{\frac{n}{2} \cdots (\frac{n}{2}-j+1) \cdot (2j) \cdots (j+1)}{n \cdots (n-2j+1)} \\
&\leq (\frac{1}{2})^j \cdot (\frac{2j}{n-j})^j \;=\; (\frac{j}{n-j})^j \;\leq\; (\frac{4j}{3n})^j. \quad (9)
\end{aligned}
$$

The last inequality follows from the fact $j \leq n/4$. We next bound $|\,\mathrm{sp}_{2j}(A)|$ for small values of j:

PROPOSITION 5. *For every $1 \leq j \leq \lfloor c/2 \rfloor$, $|\,\mathrm{sp}_{2j}(A)| \leq \sum_{i=1}^{2j} \binom{c}{i}$.*

PROOF. Using just the elementary row operations of Gaussian Elimination, we can transform A into a matrix A', which has exactly the same span as A, and that has the $c \times c$ identity matrix as a submatrix. (Recall that A has a full row rank.) It follows that any linear combination of t rows of A' results in a vector of Hamming weight at least t. Therefore, the only linear combinations to give vectors in $\mathrm{sp}_{2j}(A)$ are ones that use at most $2j$ rows of A'. The proposition follows, since the number of the latter is $\sum_{i=1}^{2j} \binom{c}{i}$. $\square$

We conclude that for $1 \leq j \leq \lfloor c/2 \rfloor$, $|\,\mathrm{sp}_{2j}(A)| \leq \sum_{i=1}^{2j} c^i = \frac{c^{2j}-1}{c-1} \cdot c \leq 2c^{2j}$ (assuming $c \geq 2$). On the other hand, we have for all $1 \leq j \leq n/4$, $|\,\mathrm{sp}_{2j}(A)| \leq |\,\mathrm{sp}(A)| \leq 2^c$. Note that the quantity $2c^{2j}$ exceeds 2^c, when $j \geq \frac{c-1}{2 \log c}$. We thus define $\ell \stackrel{\mathrm{def}}{=} \lfloor \frac{c-1}{2 \log c} \rfloor$ and break the sum on the RHS of Equation 8, which we need to bound, into two parts as follows:

$$
\begin{aligned}
&\sum_{j=1}^{n/4} |\,\mathrm{sp}_{2j}(A)| \cdot \frac{\binom{\frac{n}{2}}{j}}{\binom{n}{2j}} \\
&\quad= \sum_{j=1}^{\ell} |\,\mathrm{sp}_{2j}(A)| \cdot \frac{\binom{\frac{n}{2}}{j}}{\binom{n}{2j}} + \sum_{j=\ell+1}^{n/4} |\,\mathrm{sp}_{2j}(A)| \cdot \frac{\binom{\frac{n}{2}}{j}}{\binom{n}{2j}} \\
&\quad\leq \sum_{j=1}^{\ell} (2c^{2j}) \cdot (\frac{4j}{3n})^j + 2^c \cdot \max_{\ell < j \leq n/4} (\frac{4j}{3n})^j. \qquad (10)
\end{aligned}
$$

The last inequality follows from Equation 9, from Proposition 5, and from the fact $\sum_{j=\ell+1}^{n/4} |\,\mathrm{sp}_{2j}(A)| \leq |\,\mathrm{sp}(A)| \leq 2^c$. We bound each of the terms on the RHS of Equation 10 separately. We start with the first one:

$$
\sum_{j=1}^{\ell} (2c^{2j}) \cdot (\frac{4j}{3n})^j \;=\; 2 \cdot \sum_{j=1}^{\ell} (\frac{4c^2 j}{3n})^j \;\leq\; 2 \cdot \sum_{j=1}^{\ell} (\frac{4c^2 \ell}{3n})^j
$$

Recall that we assumed $c^3/\log c \leq 3n/4$. Hence, $4c^2\ell/(3n) \leq 2c^3/(3n \log c) \leq 1/2$. We can thus bound the geometric series as follows:

$$
\begin{aligned}
2 \cdot \sum_{j=1}^{\ell} (\frac{4c^2\ell}{3n})^j &\leq 2 \cdot \frac{4c^2\ell}{3n} \cdot \frac{1}{1 - \frac{4c^2\ell}{3n}} \;\leq\; \frac{16c^2\ell}{3n} \\
&\leq \frac{8c^3}{3n \log c}. \qquad (11)
\end{aligned}
$$

We now turn to bounding the second term on the RHS of Equation 10.

PROPOSITION 6. *The function $g(j) = (aj)^j$, where $a > 0$, has a local minimum at $j^* = \frac{1}{ae}$ in the interval $(0, \infty)$.*

PROOF. We rewrite g as follows: $g(j) = e^{j \ln(aj)}$. The derivative of g is the following:

$$
g'(j) \;=\; e^{j \ln(aj)} \cdot (\ln(aj) + 1).
$$

Thus, g has a local extremum at $j^* = \frac{1}{ae}$. We next verify it is a local minimum. The second derivative of g is the following:

$$
g''(j) \;=\; g'(j) \cdot (\ln(aj)+1) + g(j) \cdot \frac{1}{j} \;=\; g(j) \cdot ((\ln(aj)+1)^2 + \frac{1}{j}).
$$

Since g is positive in the interval $(0, \infty)$, then $g''(j) > 0$ for all j in this interval. In particular, $g''(j^*) > 0$, implying j^* is a local minimum. $\square$

Proposition 6 shows that the function $g(j) = (aj)^j$ has a local minimum at $j^* = \frac{1}{ae}$ in the interval $(0, \infty)$. In our case $a = \frac{4}{3n}$, and thus $j^* = 3n/(4e) \geq n/4$. Therefore the maximum of $(\frac{4j}{3n})^j$ in the interval $[\ell, n/4]$ is obtained at $j = \ell$. We conclude that:

$$
\begin{aligned}
2^c \cdot \max_{\ell < j \leq n/4} (\frac{4j}{3n})^j &\leq 2^c \cdot (\frac{4\ell}{3n})^\ell \;\leq\; 2^c \cdot (\frac{2c}{3n \log c})^{\frac{c}{2 \log c}} \\
&\leq (\frac{4c}{3n \log c})^c \;\leq\; (\frac{2c}{n})^c \;\leq\; \frac{c}{n} \\
&\leq \frac{c^3}{n \log c}. \qquad (12)
\end{aligned}
$$

In the next to the last inequality we used the fact $2 \leq c \leq n/4$. Combining Equations 10, 11, and 12, we have

$$
\sum_{j=1}^{n/4} |\,\mathrm{sp}_{2j}(A)| \cdot \frac{\binom{\frac{n}{2}}{j}}{\binom{n}{2j}} \;\leq\; \frac{8c^3}{3n \log c} + \frac{c^3}{n \log c} \;\leq\; O(\frac{c^3}{n \log c}).
$$

This completes the proof of Lemma 2.

8. OPEN PROBLEMS

The main question in quantum communication complexity is to characterize its power in relation to classical communication complexity. For partial Boolean functions it was known that quantum two-way communication complexity could be exponentially lower than the classical one [16]. Here we prove a similar result even for one-way communication complexity. The main open question is what is the relationship between quantum and classical communication complexity for total functions. Are they polynomially related for all total functions? Is this relationship even tighter in the case of one-way communication complexity? Moreover, can we show an exponential separation between quantum one-way communication complexity and randomized two-way communication complexity?

It is also very intriguing to study the connection between quantum one-way communication complexity and quantum advice and proofs. For example, can our result be used to prove an oracle separation between the classes *BQP/poly* and *BQP/qpoly* or between *QMA* and *QCMA*?

Acknowledgements

We would like to thank Umesh Vazirani, Ashwin Nayak and Kunal Talwar for helpful discussions.

9. REFERENCES

[1] S. Aaronson. Limitations of quantum advice and one-way communication. In *Proceedings of the 19th IEEE Conference on Computational Complexity (CCC)*, 2004. to appear.

[2] A. Ambainis, L. J. Schulman, A. Ta-Shma, U. V. Vazirani, and A. Wigderson. The quantum communication complexity of sampling. *SIAM J. on Computing*, 32(6):1570–1585, 2003.

[3] L. Babai and P. G. Kimmel. Randomized simultaneous messages: Solution of a problem of Yao in communication complexity. In *Proceedings of the 12th IEEE Conference on Computational Complexity (CCC)*, pages 239–246, 1997.

[4] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16), 2001.

[5] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the 30th ACM Symposium on Theory of Computing (STOC)*, pages 63–68, 1998.

[6] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[7] P. Frankl and R. M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357–368, 1981.

[8] M. Karchmer and A. Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM J. on Computing*, 26(5):255–265, 1990.

[9] I. Kerenidis and R. de Wolf. Exponential lower bound for 2-query locally decodable codes via quantum argument. In *Proceedings of the 35th ACM Symposium on Theory of Computing (STOC)*, pages 106–115, 2003.

[10] H. Klauck. On quantum and probabilistic communication: Las Vegas and one-way protocols. In *Proceedings of the 32nd ACM Symposium on Theory of Computing (STOC)*, pages 644–651, 2000.

[11] I. Kremer. *Quantum Communication*. Master's Thesis, The Hebrew University of Jerusalem, 1995.

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

[13] I. Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39:67–71, 1991.

[14] I. Newman and M. Szegedy. Public vs. private coin flips in one round communication games. In *Proceedings of the 28th ACM Symposium on Theory of Computing (STOC)*, pages 561–570, 1996.

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

[16] R. Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the 31st ACM Symposium on Theory of Computing (STOC)*, pages 358–367, 1999.

[17] P. Sen and S. Venkatesh. Lower bounds in the quantum cell probe model. In *Proceedings of the 28th ICALP, volume 2076*, pages 358–369, 2001.

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

[19] A. C.-C. Yao. Some complexity questions related to distributive computing. In *Proceedings of the 11th ACM Symposium on Theory of Computing (STOC)*, pages 209–213, 1979.

[20] A. C.-C. Yao. Lower bounds by probabilistic arguments. In *Proceedings of the 24th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 420–428, 1983.

[21] A. C.-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 352–361, Los Alamitos, CA, 1993.