A lower bound for the bounded round quantum communication complexity of set disjointness

Rahul Jain* Jaikumar Radhakrishnan† Pranab Sen‡

Abstract

We show lower bounds in the multi-party quantum communication complexity model. In this model, there are t parties where the ith party has input $X_i \subseteq [n]$. These parties communicate with each other by transmitting qubits to determine with high probability the value of some function F of their combined input $(X_1, \ldots, X_t)$. We consider the class of boolean valued functions whose value depends only on $X_1 \cap \cdots \cap X_t$; that is, for each F in this class there is an $f_F : 2^{[n]} \to \{0,1\}$, such that $F(X_1, \ldots, X_t) = f_F(X_1 \cap \cdots \cap X_t)$. We show that the t-party k-round communication complexity of F is $\Omega(s_m(f_F)/(k^2))$, where $s_m(f_F)$ stands for the 'monotone sensitivity of f_F' and is defined by $s_m(f_F) \stackrel{\Delta}{=} \max_{S \subseteq [n]} |\{i : f_F(S \cup \{i\}) \neq f_F(S)\}|$.

For two-party quantum communication protocols for the set disjointness problem, this implies that the two parties must exchange $\Omega(n/k^2)$ qubits. An upper bound of $O(n/k)$ can be derived from the $O(\sqrt{n})$ upper bound due to Aaronson and Ambainis [AA03]. For $k = 1$, our lower bound matches the $\Omega(n)$ lower bound observed by Buhrman and de Wolf [BdW01] (based on a result of Nayak [Nay99]), and for $2 \leq k \ll n^{1/4}$, improves the lower bound of $\Omega(\sqrt{n})$ shown by Razborov [Raz02]. For protocols with no restrictions on the number of rounds, we can conclude that the two parties must exchange $\Omega(n^{1/3})$ qubits. This, however, falls short of the optimal $\Omega(\sqrt{n})$ lower bound shown by Razborov [Raz02].

Our result is obtained by adapting to the quantum setting the elegant information-theoretic *arguments of Bar-Yossef, Jayram, Kumar and Sivakumar [BJKS02b]. Using this method we can show similar lower bounds for the $\mathcal{L}_\infty$ function considered in [BJKS02b].*

1 Introduction

Classical communication complexity: The (classical) communication complexity model of Yao [Yao79] provides an abstract setting for studying the classical communication required for computing a function whose inputs are distributed between several parties. In its most widely studied version, there are two parties, Alice and Bob with inputs $X_A, X_B \subseteq [n]$, who exchange classical messages based on a fixed protocol in order to determine the value of some function $F(X_A, X_B)$. The goal is to design a protocol so that the parties need to exchange as few bits as possible. This model of communication is relatively well-understood (see the book of Kushilevitz and Nisan [KN97]) both in the deterministic and the randomised setting. In this paper, we will be interested in the randomised setting, where the parties are allowed to err with some small probability (say at most $\frac{1}{3}$). Tight lower bounds are known for several functions in this model, for example, the equality function $X_A \stackrel{?}{=} X_B$ [Yao79, LS81], the set disjointness function DISJ $X_A \cap X_B \stackrel{?}{=} \emptyset$ [KS92, Raz92] and the inner-product function $|X_A \cap X_B| \stackrel{?}{=} 0 \bmod 2$ [CG88].

Quantum communication complexity: In [Yao93], Yao introduced the two-party quantum communication model in order to investigate if communication costs for computing functions distributively reduces significantly when the parties are allowed to exchange qubits and perform quantum operations locally. Since then, there has been a flurry of results in this model. We will be mainly interested in the bounded error version of this model, where the two parties are allowed to err with some small probability (say at most $\frac{1}{3}$). It was observed early that for the equality and the inner-product functions the quantum model does not provide any significant savings: the complexity of the equality function is still $\Theta(\log n)$ [Kre95] and the complexity of the inner-product function is still $\Theta(n)$ [Kre95, CvDNT98].

*School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India and CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands. Email: rahulj@tcs.tifr.res.in. Partially supported by the Kanwal Rekhi Career Development Scholarship, and by the EU fifth program grant RESQ and NWO grant 612.055.001.

†School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India. Email: jaikumar@tcs.tifr.res.in.

‡Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1, Canada. Email: p2sen@iqc.ca. Work done while the author was visiting TIFR, Mumbai.

The set disjointness function: For the set disjointness function, however, quantum protocols were found to be strictly more powerful than their classical randomised counterparts. Since the communication complexity of the set disjointness function is central to the work presented in this paper, we describe its history in greater detail. In the bounded error classical setting Babai, Frankl and Simon [BFS86] showed a lower bound of $\Omega(\sqrt{n})$. This was improved to an $\Omega(n)$ lower bound by Kalyanasundaram and Schnitger [KS92]; their proof was simplified by Razborov [Raz92]. There is a straightforward protocol with $n+1$ bits of communication where Alice sends her entire input to Bob, who computes the answer and returns it to Alice. Interest in the communication complexity of several problems related to the set disjointness function has been revived recently because of their connection to showing lower bounds in the classical data stream model [AMS99, FKS02, GGI$^+$02, Ind00, GMMO00, JKS03, SS02]. One of these problems is the $\mathcal{L}_\infty$ promise problem: Alice and Bob are given inputs $X_A, X_B \in \{0,1,\ldots,m\}^n$, with the promise that either for all $i \in [n]$, $|X_A[i]-X_B[i]| \leq 1$ or there exists an $i \in [n]$, such that $|X[i]-Y[i]| = m$; they must communicate in order to distinguish between these two types of inputs. For this problem, Saks and Sun [SS02] showed a lower bound of $\Omega(n/m^2)$ in a restricted model; their lower bound was strengthened by Bar-Yossef, Jayram, Kumar and Sivakumar [BJKS02b], who obtained the same lower bound without any restrictions.

The quantum communication complexity of set disjointness was first studied by Buhrman, Cleve and Wigderson [BCW98], who showed that there is a protocol for this problem with $O(\sqrt{n}\log n)$ qubits of communication. This bound was improved to $O(\sqrt{n}c^{\log^* n})$, where c is a small constant, by Hoyer and de Wolf [HdW02], and recently to $O(\sqrt{n})$ by Aaronson and Ambainis [AA03]. By a result of Razborov [Raz02] this last bound is optimal.

Multi-party classical communication complexity: In fact, there are several ways to generalise the two-party model to the multi-party model. In this paper, we will consider the version where there are t parties $P_1, P_2, \ldots, P_t$ with respective inputs $X_1, X_2, \ldots, X_t \subseteq [n]$. In each round of communication some party sends a message to another party. The party who receives the last message can determine the desired value $F(X_1, X_2, \ldots, X_t)$ based on his current state at that point. Recently, because of its connection to the problem of computing *frequency moments* in the data stream model [AMS99], the following *promise set disjointness* problem has been studied. Here, the parties are required to distinguish between two types of inputs: in the first type, $X_1, X_2, \ldots, X_t$ are pairwise disjoint; in the second type, $X_1, X_2, \ldots, X_t$ have exactly one element in common but are otherwise pairwise disjoint. For this problem, Chakrabarti, Khot and Sun [CKS03] show a lower bound of $\Omega(n/(t\log t))$, improving an earlier $\Omega(n/t^2)$ lower bound of Bar-Yossef, Jayram, Kumar and Sivakumar [BJKS02b] and an $\Omega(n/t^4)$ lower bound of Alon, Matias and Szegedy [AMS99]. A slight variant of this problem, called the approximate set disjointness problem, was considered by Nisan [Nis02]; the lower bounds mentioned above apply to Nisan's version as well. The multi-party quantum communication complexity of these problems has not been considered before this work.

1.1 Our results

The upper and lower bounds on the two-party quantum communication complexity of the set disjointness function are tight up to constant factors, if there are no restrictions imposed on the number of rounds (i.e. the number of messages) in the protocol. The best upper bound uses $O(\sqrt{n})$ rounds of communication, and from it one can derive a k-round protocol where the parties exchange a total of at most $O(n/k)$ qubits. For $k=1$, Buhrman and de Wolf [BdW01] observed that a lower bound of $\Omega(n)$ follows from the results of Nayak [Nay99] for the index-function problem. For $k \geq 2$, Klauck, Nayak, Ta-Shma and Zuckerman [KNTZ01] showed a lower bound of $\Omega(n^{1/k})$, but this is subsumed by Razborov's [Raz02] lower bound of $\Omega(\sqrt{n})$ which holds even if there is no restriction on the number of rounds. However, for small k, Razborov's lower bound is far from the best upper bound known, namely $O(n/k)$. Our first result implies lower bounds for the two-party bounded error k-round quantum communication complexity of set disjointness that comes closer to the upper bound of $O(n/k)$. In fact, the result holds for a multi-party quantum communication model which we define in detail in Section 2.2. In this model, there are t parties where the ith party has input $X_i \subseteq [n]$. These parties communicate with each other by transmitting qubits from one party to another to determine with high probability the value of some function F of their combined input $(X_1, \ldots, X_t)$. We consider the class of boolean valued functions whose value depends only on $X_1 \cap \cdots \cap X_t$; that is, for each F in this class there is an $f_F : 2^{[n]} \to \{0,1\}$, such that $F(X_1, \ldots, X_t) = f_F(X_1 \cap \cdots \cap X_t)$. We call such functions F *set disjointness-like*. Define the 'monotone sensitivity of f_F' as $s_m(f_F) \stackrel{\Delta}{=} \max_{S \subseteq [n]} |\{i : f_F(S \cup \{i\}) \neq f_F(S)\}|$.

Result 1 *The t-party k-round bounded error quantum communication complexity of a set disjointness-like function F is $\Omega(s_m(f_F)/k^2)$.*

In fact, Result 1 follows from the following result via easy reductions.

Result 1' *The t-party k-round bounded error quantum communication complexity of the promise set disjointness problem is $\Omega(n/k^2)$. This lower bound also holds for Nisan's approximate set disjointness problem [Nis02].*

Remarks:
1. Observe that the lower bound in Result 1' is independent of t! This appears to contradict the $O((n \log n)/t)$ upper bound for the promise set disjointness problem in [BJKS02a]. However, that upper bound is in the multi-party *simultaneous message* model, whereas in our definition of multi-party quantum protocols it is required to pass messages from one party to another. Thus, the simultaneous message protocol of [BJKS02a] is actually a t-round protocol in our model.
2. For two-party quantum protocols with an unbounded number of rounds, we get a lower bound of $\Omega(n^{1/3})$ for the set disjointness problem.

For the $\mathcal{L}_\infty$ promise problem we get the following lower bound.

Result 2 *The two-party k-round quantum communication complexity of the $\mathcal{L}_\infty$ promise problem is $\Omega(n/(k^3 m^{k+1}))$.*

All our lower bounds hold even if the parties start with arbitrary prior entanglement that is independent of the inputs.

Finally, we remark that our quantum communication complexity lower bounds imply space lower bounds for a natural model of 'quantum data stream computation', in exactly the same way as in the classical setting.

1.2 Techniques used

The original lower bounds for set disjointness in the classical two-party communication model are based on deep analyses of the communication matrix and can be said to be based on the *discrepancy method* (see e.g. [KN97]). Razborov's recent $\Omega(\sqrt{n})$ lower bound [Raz02] for the bounded error two-party quantum communication complexity of set disjointness also uses the discrepancy method. The discrepancy method for quantum protocols was formulated explicitly by Kremer [Kre95] (see also Klauck [Kla01] and Yao [Yao93]), but Razborov's proof extends it substantially by developing interesting and powerful tools based on the spectral theory of matrices.

Recently however, Bar-Yossef et al. [BJKS02b] proposed an information-theoretic approach for studying set disjointness-like problems in the classical setting. Using a refinement of the notion of *information cost* of a communication protocol originally defined by Chakrabarti, Shi, Wirth and Yao [CSWY01], they showed that a linear lower bound for the bounded error two-party randomised communication complexity of set disjointness follows from an $\Omega(1)$ lower bound on a certain *information cost* of a two-party communication protocol computing the AND $a \wedge b$ of just two bits a, b! The information-theoretic machinery essentially allowed them to treat the set disjointness function like a direct sum of n two-bit AND 's. Their work provided a compelling and beautiful illustration of information-theoretic tools in the analysis of communication protocols. Interestingly, the idea of proving lower bounds for set disjointness by treating it like a direct sum of n two-bit AND 's was earlier employed in [KKN95] in the setting of two-party nondeterministic classical communication complexity; however, their approach was not information-theoretic and does not seem to be suitable for bounded error classical randomised or quantum communication protocols.

We adapt their approach to the quantum setting. In order to bring out the contribution of this paper more clearly, we will now informally describe the information-theoretic argument underlying the proof of [BJKS02b] and discuss how we adapt it to the quantum setting. The argument has two parts: in the first part, using a direct-sum property for information cost of a communication protocol one reduces the communication problem DISJ to the communication problem AND of two bits (one with Alice and one with Bob); in the second part, one shows that any communication protocol for AND of two bits needs to have high information cost.

The information cost approach: The first part of the argument is based on the notion of *information cost* of private coin randomised communication protocols, defined to be the *Shannon mutual information* between the inputs (which are assumed to come from some distribution) and the entire message transcript of the protocol. Bar-Yossef et al. [BJKS02b] examine the information cost of the protocol for several distributions. Let the number of bits transmitted by the protocol be c. Then, the information cost is also bounded by c for each distribution.

At this point it will be convenient to view the inputs X_A and X_B of Alice and Bob as elements of $\{0,1\}^n$ and the set disjointness function DISJ as $\bigvee_{i=1}^n X_A[i] \wedge X_B[i]$. A typical distribution considered by Bar-Yossef et al. is defined as follows. For each coordinate i, independently, one party is given the input 0 and the other party is given a uniformly random bit. Using the sub-additivity property of mutual information, one concludes that the sum over i of the mutual information between the transcript and $X_A[i]$ is bounded by c; a similar statement holds for Bob's inputs. It is then not hard to argue using a standard averaging argument that there is an i and a probability distribution D^* on $(X_A[j], X_B[j] : j \neq i)$ such that the following conditions

hold:

- $X_A[j], X_B[k], j \neq i, k \neq i$ are independent random variables under D^*;
- For all $j \neq i$, $X_A[j] \wedge X_B[j] = 0$ (with probability 1);
- If $X_A[i]$ is set to 0, $X_B[i]$ is chosen uniformly at random from $\{0,1\}$ and $(X_A[j], X_B[j] : j \neq i)$ are chosen according to D^*, then the mutual information between the message transcript and $X_B[i]$ is at most $2c/n$; similarly, if $X_B[i]$ is set to 0, $X_A[i]$ is chosen uniformly at random from $\{0,1\}$ and $(X_A[j], X_B[j] : j \neq i)$ are chosen according to D^*, then the mutual information between the message transcript and $X_A[i]$ is at most $2c/n$.

From the first condition, by viewing $(X_A[j], X_B[j] : j \neq i)$ as private coins of the two parties, we obtain from the protocol for DISJ a protocol that computes the AND of the two bits $X_A[i]$ and $X_B[i]$. The stage is thus set for analysing the information cost of a protocol computing the AND of two bits: a lower bound of ϵ on this quantity translates to a lower bound of $\Omega(\epsilon n)$ on the communication complexity of the set disjointness function.

In order to implement this programme in the quantum setting, one has to define a notion of information cost for quantum protocols. It is not immediately clear how this can be done, because quantum operations are notorious for destroying the states on which they act; in particular, it is not reasonable to expect that the complete transcript of all messages is part of the final global state of the algorithm. Even if the complete transcript is available in the final global state of the algorithm, it may not contain any information about the inputs of either party. If the parties are allowed prior entanglement, then using quantum teleportation, one can implement any protocol such that the messages are classical and uniformly random. So, the transcript will just be a uniformly random string of length c independent of the actual inputs!

The definition of information loss for quantum protocols: We address these difficulties as follows. Assume that the players' inputs come from some classical probability distribution. Without loss of generality, the players make a 'safe' copy of their (classical) inputs before proceeding with the quantum protocol. Instead of considering the information carried by a particular message, we examine the the context in which the message is received i.e. we consider the von Neumann mutual information between the sender's input and all the qubits in the possession of the receiver at that time, including the qubits of the message just received. The *information loss* (we use the term loss instead of cost) of the protocol for the given input distribution is defined to be a certain weighted sum of these mutual informations taken over all rounds. With this definition of information loss, the arguments of [BJKS02b] are easily carried over to the quantum setting. We can then conclude that if the information loss of computing the AND of two bits is ϵ then the communication complexity of DISJ is $\Omega(n\epsilon/k)$.

We have arrived at the second part of our programme, that is, to show non-trivial lower bounds on the information loss of a quantum protocol computing the AND of two bits. In their original argument, [BJKS02b] showed a lower bound on the information cost of a classical private coin protocol computing the AND of two bits via a direct argument using *Hellinger distances* between certain probability distributions. Since we are working with our different notion of information loss, this argument does not appear to be immediately applicable to us; so instead of reviewing it, we will now directly describe our new argument for showing a lower bound on the information loss of a quantum protocol computing the AND of two bits. We consider two input distributions: in the first distribution, Alice has 0 and Bob has a uniformly random bit; in the second distribution, Bob has 0 and Alice has a uniformly random bit. Suppose we are given that for these distributions at no stage do the qubits of the receiver of a message contain more than ϵ bits of information about the input of the sender. We wish to show that if ϵ is very small, then this leads to a contradiction. Our argument can be understood at an intuitive level in the framework of *round elimination* in communication protocols [MNSW98, KNTZ01, Sen03]. Suppose Alice sends the first message of the protocol. We know that when Bob's input is 0 the state of his qubits after receiving the first message is essentially the same whether Alice's input is 0 or 1. So no matter what her actual input is, Alice might as well send her first message assuming that her input is 0. Using standard arguments (see below), we can eliminate the first message of Alice and obtain a protocol with one fewer round of communication, increasing the error probability of the protocol by a small amount. Now it is Bob's turn. Our hypothesis says that when Alice's input is 0 the state of her qubits after receiving the first message from Bob is essentially the same whether Bob's input is 0 or 1. But the modified protocol so far has proceeded as if Alice's input is 0 (even though her actual input might be something else). We can thus eliminate Bob's first message as well. If ϵ is small, then the increase in error probability on account of this manoeuvre is also small. Proceeding in this manner we eliminate all rounds. But it is obvious that if the parties exchange no messages they cannot compute any non-trivial function unless one allows error probability greater than $1/2$. Since there are at most k rounds of communication, this gives us a lower bound of the form $\epsilon \geq \epsilon(k)$. Using these ideas one can show an $\Omega(n/k^2)$ lower bound on the two-party quantum communication complexity of the set disjointness function.

There are two aspects of our proof that require further comment.

Local transition: Recall the argument used above to eliminate Alice's first message. We know that when Bob's input is 0, the state of his qubits after receiving the first message is roughly the same whether Alice's input is 0 or 1. her input is 0. However, this does not immediately imply that the error probability of the modified protocol is not changed by much. The final answer is not just a function of Bob's state but the combined state of Alice and Bob. In particular, even though Bob's state is similar after the first round for the two inputs of Alice, his work qubits might be entangled with Alice's qubits differently in the two cases. This problem arises often in round elimination arguments and by now standard solutions exist for it by considering the *fidelity* between quantum states. This allows Alice to perform a *local transition* [KNTZ01] on her work qubits, in order to restore them to the correct state should she discover later that her actual input is 1 (recall that in the modified protocol, Alice prepares her first message assuming that her input is always 0). We use a stronger local transition lemma (Lemma 1) than the one in [KNTZ01]. The stronger lemma is crucial for getting an $\Omega(n/k^2)$ lower bound in Result 1; the local transition lemma of [KNTZ01] gives an $\Omega(n/k^4)$ lower bound.

A paradox?: In our notion of information loss of quantum protocols it is important that the parties start in a *pure* global state. In fact, this notion is unsuited for classical private coin randomised communication complexity. Consider the following classical private coin protocol for computing the AND of two bits (a, b). Alice sends Bob a random bit r, retaining a copy of r if and only if $a = 1$. Bob sends Alice $r \oplus b$; if $a = 1$, Alice can recover b using the copy of r she has and determine $a \wedge b$. Now clearly, when Bob's input is 0 he has no information about Alice's input at the end of the first round; also when Alice's input is 0 she has no information about Bob's input at the end of the second round because she does not retain a copy of r in this case. So, according to our definition this protocol has zero information loss for both the distributions considered above. Yet, the protocol computes the AND of two bits correctly! Interestingly, no such quantum protocol starting with a pure global state is possible.

1.3 The rest of the paper

In the next section, we give some definitions used in the rest of the paper. In Section 3, we prove Result 1'. The proof of Result 2 is omitted from this extended abstract.

2 Preliminaries

2.1 Information theoretic background

We now state some basic facts from information theory that we need. For a good account of quantum information theory, see e.g. [NC00].

In this paper, all quantum systems are finite dimensional and all classical random variables have finite range. Suppose A, B, C are three disjoint quantum systems having some joint density matrix ρ. Let ρ_A be the reduced density matrix of A. Then $S(A) \triangleq S(\rho_A) \triangleq -\text{Tr}\,\rho_A \log \rho_A$ is the *von Neumann entropy* of A. The *von Neumann mutual information* of A and B is defined as $I(A : B) \triangleq S(A) + S(B) - S(AB)$. The *conditional von Neumann mutual information* of A and B given C is defined as $I((A : B) \mid C) \triangleq S(AC) + S(BC) - S(C) - S(ABC)$. If C is a classical random variable taking the classical value $|c\rangle$ with probability p_c, it is easy to see that $I((A : B) \mid C) = \sum_c p_c I(A^c : B^c)$, where $(AB)^c$ denotes the joint density matrix of A and B when $C = |c\rangle$. We also write $I((A : B) \mid C = c)$ for $I(A^c : B^c)$. Mutual information satisfies the following *monotonicity* property: $I(A : BC) \geq I(A : B)$.

Suppose $D, X_1, \ldots, X_t$ are classical random variables. We say that D *partitions* $X \triangleq (X_1, \ldots, X_t)$ if for all possible values d that D can take, $X_1, \ldots, X_t$ are independent conditioned on the event $D = d$.

Fact 1 (Sub-additivity) *Let $X_1, \ldots, X_n$ be independent classical random variables. Let M be a quantum encoding of $X \triangleq (X_1, \ldots, X_n)$. Then, $I(X : M) \geq \sum_{i=1}^n I(X_i : M)$.*

For classical probability distributions P, Q on the same sample space Ω, their total variation distance is defined as $\|P - Q\|_1 \triangleq \sum_{\omega \in \Omega} |P(\omega) - Q(\omega)|$. For density matrices ρ, σ over the same Hilbert space, their trace distance is defined as follows: $\|\rho - \sigma\|_t \triangleq \text{Tr}\,\sqrt{(\rho - \sigma)^\dagger (\rho - \sigma)}$. The importance of trace distance as a metric on density matrices stems from the following fact.

Fact 2 (see [AKN98]) *Let ρ, σ be density matrices in the same finite dimensional Hilbert space $\mathcal{H}$. Let $\mathcal{F}$ be a measurement (POVM) on $\mathcal{H}$. Then, $\|\mathcal{F}\rho - \mathcal{F}\sigma\|_1 \leq \|\rho - \sigma\|_t$.*

Suppose A, B are disjoint quantum systems. Let ρ_{AB}, σ_{AB} be two density matrices of the joint quantum system AB. The trace distance satisfies the following property of *monotonicity*: $\|\rho_{AB} - \sigma_{AB}\|_t \geq \|\rho_A - \sigma_A\|_t$. In fact, Fact 2 can be derived from the monotonicity of trace distance.

We now state our strong local transition lemma. The proof uses an inequality implicitly contained in [Lin91] (see

also [FvdG99, DHR78]). It is omitted from this extended abstract.

Lemma 1 *Suppose X and Q are disjoint quantum systems, where X is a classical random variable uniformly distributed over $\{0,1\}$ and Q is a quantum encoding $x \to \sigma_x$ of X. Let $\mathcal{H}$ denote the Hilbert space of Q. Let $\mathcal{K}$ be any Hilbert space of dimension at least the dimension of $\mathcal{H}$, and $|\phi_0\rangle, |\phi_1\rangle$ any purifications of σ_0, σ_1 respectively in $\mathcal{H} \otimes \mathcal{K}$. Then there is a local unitary transformation U on $\mathcal{K}$ that maps $|\phi_2\rangle$ to $|\phi_2'\rangle \triangleq (I \otimes U)|\phi_2\rangle$ (I is the identity operator on $\mathcal{H}$) such that $\||\phi_1\rangle\langle\phi_1| - |\phi_2'\rangle\langle\phi_2'|\|_t \leq \sqrt{8I(X:Q)}$.*

2.2 Quantum communication complexity

We define t-party quantum communication protocols as a natural extension of two-party quantum communication protocols defined by Yao [Yao93]. Let $f : \mathcal{X}_1 \times \mathcal{X}_2 \cdots \mathcal{X}_t \to \mathcal{Z}$ be a function. There are t parties $\mathcal{P}_1, \ldots, \mathcal{P}_t$ who hold qubits. When the quantum communication protocol Π starts, $\mathcal{P}_i$ holds $|x_i\rangle$ where $x_i \in \mathcal{X}_i$ together with some ancilla qubits ('work qubits') in the state $|0\rangle$. $\mathcal{P}_1, \ldots, \mathcal{P}_t$ may also share an input independent prior entanglement pure state (say $|\psi\rangle$). Different parties possess different qubits of $|\psi\rangle$. The parties take turns to communicate to compute $f(x_1, x_2, \cdots, x_t)$. Suppose it is $\mathcal{P}_1$'s turn to communicate. $\mathcal{P}_1$ can make an arbitrary unitary transformation on the qubits in her possession at this time and then send some of her qubits to $\mathcal{P}_2, \ldots, \mathcal{P}_t$. Whose turn it is to communicate, the unitary transformation applied by the active player and the qubits that the active player sends to the other players are predetermined by Π and independent of the input $(x_1, \ldots, x_t)$. A *round* of communication denotes the qubits that the active player sends to the other players. Sending qubits does not change the overall superposition, but rather changes the ownership of the qubits. At the end of the protocol Π, one of the parties performs a von Neumann measurement in the computational basis of some qubits in her possession (the 'answer qubits') to output an answer $\Pi(x_1, x_2, \cdots, x_t)$. The party performing the measurement as well as the qubits that she measures are predetermined by Π and independent of the input $(x_1, \ldots, x_t)$. We say that protocol Π computes f with error δ if $\max_{x_1,\ldots,x_t} \Pr[\Pi(x_1, \ldots, x_t) \neq f(x_1, \ldots, x_t)] \leq \delta$. The communication cost of Π is the number of qubits exchanged in Π between all the parties. The t-party k-round δ-error quantum communication complexity of f, denoted by $Q_\delta^{t,k}(f)$, is the minimum communication cost of a t-party k-round δ-error quantum protocol with prior entanglement for f. When δ is omitted, we mean that $\delta = 1/3$.

We require that the parties make a 'safe' copy of their inputs (using for example CNOT gates) before beginning the protocol Π. This is possible without loss of generality because the inputs are in computational basis states. Thus, the input qubits of the parties are never sent as messages, their state remains unchanged throughout the execution of Π, and they are never measured i.e. some work qubits are measured to determine the result $\Pi(x_1, \ldots, x_t)$. We call such protocols *safe*, and henceforth, we will assume that all our protocols are safe.

Fact 3 (see [CvDNT98]) *Let* Alice *have a classical random variable X. Suppose* Alice *and* Bob *share a pure state on some qubits (prior entanglement) independent of X. Initially* Bob*'s qubits have no information about X. Now let* Alice *and* Bob *run a quantum communication protocol, at the end of which* Bob*'s qubits possess m bits of information about X. Then,* Alice *has to totally send at least $m/2$ qubits to* Bob*.*

We now define the *conditional information loss* of a t-party quantum communication protocol with prior entanglement. For technical reasons, we need to work with a *conditional* version of information loss instead of the unconditional version described in the introduction. A similar *conditional* version of information cost is used in [BJKS02b] to prove their lower bounds. But first, we need a couple of preliminary definitions.

Definition 1 (Embedding) *For $\mathbf{x} \in \mathcal{X}^n$, $j \in [n]$, and $x \in \mathcal{X}$, let $\mathsf{embed}(\mathbf{x}, j, x)$ be the element of $\mathcal{X}^n$ obtained by replacing $\mathbf{x}[j]$ by x, that is, $\mathsf{embed}(\mathbf{x}, j, x)[\ell] \triangleq \mathbf{x}[\ell]$ for $\ell \neq j$, and $\mathsf{embed}(\mathbf{x}, j, x)[j] \triangleq x$.*

Definition 2 (Collapsing) *Suppose $F : \mathcal{X}^n \to \mathcal{Z}$. We say that $\mathbf{x} \in \mathcal{X}^n$ collapses F to the function $h : \mathcal{X} \to \mathcal{Z}$ if for all $u \in \mathcal{X}$, $j \in [n]$, $F(\mathsf{embed}(\mathbf{x}, j, u)) = h(u)$. We say that a random variable $\mathbf{X}$ taking values in $\mathcal{X}^n$ collapses F to h if it collapses F to h with probability 1.*

Let $D, X_1, \ldots, X_t$ be classical random variables taking values in some finite sets $\mathcal{D}, \mathcal{X}_1, \ldots, \mathcal{X}_t$ respectively. Let $X \triangleq (X_1, \ldots, X_t)$. The random variable $(X, D)^n$ is obtained by taking n independent copies of (X, D). Thus, X^n takes values in $(\mathcal{X}_1 \times \cdots \times \mathcal{X}_t)^n$ which we identify with $\mathcal{X}_1^n \times \cdots \times \mathcal{X}_t^n$ in the natural way. Suppose D partitions X, and $(\mathbf{X}, \mathbf{D}) \triangleq (X, D)^n$; then it is easy to verify that $\mathbf{D}$ partitions $\mathbf{X}$. Let Π be a t-party k-round δ-error quantum protocol for computing $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_t \to \mathcal{Z}$. Suppose $X_1, \ldots, X_t$ are the random variables corresponding to the inputs of $\mathcal{P}_1, \ldots, \mathcal{P}_t$. Let $\mathcal{P}^j$ denote the active player in round j. Let X^j denote the input random variable of $\mathcal{P}^j$. $\hat{P}^j$ denote the qubits all players except $\mathcal{P}^j$ just after round j is complete. Let $k(j)$ denote the number of rounds of Π in which the player $\mathcal{P}^j$ is active.

Definition 3 (Conditional information loss) *In the notation above, the conditional information loss of Π under*

(X, D) *is defined by* $\mathsf{IL}(\Pi \mid (X, D)) \triangleq \sum_{j=1}^{k} \frac{k}{k(j)} \cdot I((X^j : \hat{P}^j) \mid D)$. *The t-party k-round δ-error conditional information loss of F under (X, D), denoted by $\mathsf{IL}_{\delta}^{t,k}(F \mid (X, D))$, is the infimum $\mathsf{IL}(\Pi \mid (X, D))$ taken over all t-party k-round δ-error quantum protocols with prior entanglement Π for F. [Note that δ upper bounds the error of Π for* all *inputs in $\mathcal{X}_1 \times \cdots \times \mathcal{X}_t$. In particular, this error bound applies even to inputs not in the support of X.]*

3 Lower bound for set disjointness

Lemma 2 *Let $F : \mathcal{X}_1^n \times \cdots \times \mathcal{X}_t^n \to \mathcal{Z}$. Let $X_1, \ldots, X_t$ be classical random variables taking values in $\mathcal{X}_1, \ldots, \mathcal{X}_t$ respectively. Define $X \triangleq (X_1, \ldots, X_t)$. Suppose X is partitioned by a classical random variable D taking values in some set $\mathcal{D}$. Let $(\mathbf{X}, \mathbf{D}) \triangleq (X, D)^n$. Suppose $\mathbf{X}$ collapses F to the function $h : \mathcal{X}_1 \times \cdots \times \mathcal{X}_t \to \mathcal{Z}$. Then, $\mathsf{IL}_{\delta}^{t,k}(h \mid (X, D)) \leq \frac{2k}{n} \cdot Q_{\delta}^{t,k}(F)$.*

Proof: Suppose Π is a t-party k-round δ-error quantum protocol with prior entanglement for F with communication cost $c \triangleq Q_{\delta}^{t,k}(F)$. Our goal is to show that there is a t-party k-round δ-error quantum protocol with prior entanglement for h having information loss at most $\frac{2kc}{n}$ under (X, D). While analysing Π, we will need to maintain that the global state of $\mathcal{P}_1, \ldots, \mathcal{P}_t$ is pure at all times. However, we will run Π on random inputs drawn from certain product probability distributions. In such a situation, we will adopt the following convention. We will assume that in addition to the usual input registers IN_i, $\mathcal{P}_i$ has another set of registers $\widetilde{\mathsf{IN}}_i$. When we require that $\mathcal{P}_i$'s inputs be some random variable $\mathbf{X}_i$, we in fact, start with the following state in the registers $\mathsf{IN}_i\widetilde{\mathsf{IN}}_i$: $\sum_{\mathbf{x} \in \mathcal{X}_i^n} \sqrt{p_{\mathbf{x}}}|\mathbf{x}\rangle|\mathbf{x}\rangle$, where $p_{\mathbf{x}} \triangleq \Pr[\mathbf{X}_i = \mathbf{x}]$. Then, we run the protocol Π as before with input registers IN_i. During this execution no quantum gates are applied to registers $\widetilde{\mathsf{IN}}_i$, they are not sent as messages and they are never measured. From now on the classical random variable $\mathbf{X}_i$ denotes the state of the registers IN_i, which stays unchanged throughout the protocol Π because Π is safe. In this revised protocol Π', $\widetilde{\mathsf{IN}}_i$ is included amongst the qubits of $\mathcal{P}_i$. Π' has the same communication cost as Π. Π' is a δ-error protocol for F with communication cost c. Consider the execution of Π' on input $\mathbf{X} \triangleq (\mathbf{X}_1, \ldots, \mathbf{X}_t)$ conditioned on $\mathbf{D} = \mathbf{d}$; note that under this condition $\mathbf{X}_1, \ldots, \mathbf{X}_t$ are independent random variables. Let $c(i)$ denote the total number of qubits sent by the party $\mathcal{P}^i$ in protocol Π' (which is the same as the total number of qubits sent by $\mathcal{P}^i$ in protocol Π). Then we have, for all $1 \leq i \leq k$, $\sum_{j=1}^{n} I((\mathbf{X}^i[j] : \hat{P}^i) \mid \mathbf{D} = \mathbf{d}) \leq I((\mathbf{X}^i : \hat{P}^i) \mid \mathbf{D} = \mathbf{d}) \leq 2c(i)$. The first inequality above follows from Fact 1 because by our definition of $(\mathbf{X}, \mathbf{D})$, $(\mathbf{X}_A[j] : 1 \leq j \leq n)$ are independent random variables when conditioned on $\mathbf{D} = \mathbf{d}$; the second inequality follows from Fact 3.

Averaging over the possible values of $\mathbf{D}$, we obtain: $\forall i, 1 \leq i \leq k, \sum_{j=1}^{n} I((\mathbf{X}^i[j] : \hat{P}^i) \mid \mathbf{D}) \leq 2c(i)$. Summing these inequalities with weight $k/k(i)$ over all rounds i, we obtain $\sum_{j=1}^{n} \sum_{i=1}^{k} \frac{k}{k(i)} \cdot I((\mathbf{X}^i[j] : \hat{P}^i) \mid \mathbf{D}) \leq 2ck$, which implies:

$$\exists j, 1 \leq j \leq n, \sum_{i=1}^{k} \frac{k}{k(i)} \cdot I((\mathbf{X}^i[j] : \hat{P}^i) \mid \mathbf{D}) \leq \frac{2ck}{n}. \quad (1)$$

Fix a value of j so that the last inequality holds. For $\mathbf{d} \in \mathcal{D}^n$, let

$$I(\mathbf{d}) \triangleq \sum_{i=1}^{k} \frac{k}{k(i)} \cdot I((\mathbf{X}^i[j] : \hat{P}^i) \mid \mathbf{D} = \mathbf{d}) \quad (2)$$

Then from (1), $\mathrm{E}_{\mathbf{D}}[I(\mathbf{D})] \leq \frac{2ck}{n}$.

We will now obtain a protocol for h by 'embedding' its input as the jth input of Π'. Using a straightforward averaging argument we first fix a value $\hat{\mathbf{d}} \in \mathcal{D}^n$ so that

$$\begin{aligned} &\sum_{d \in \mathcal{D}} \Pr[D = d] I(\mathsf{embed}(\hat{\mathbf{d}}, j, d)) \\ &\quad = \quad \mathrm{E}_D[I(\mathsf{embed}(\hat{\mathbf{d}}, j, D)) \leq \tfrac{2ck}{n}. \end{aligned} \quad (3)$$

Consider the following quantum protocol with prior entanglement Π_h for computing $h(u_1, \ldots, u_t)$. On input $u_i \in \mathcal{X}_i$, $\mathcal{P}_i$ prepares her input registers as follows. In the registers $(\mathsf{IN}_i[\ell], \widetilde{\mathsf{IN}}_i[\ell] : \ell \neq j)$ $\mathcal{P}_i$ places the superposition $\sum_{\mathbf{x} \in \mathcal{X}_i^{n-1}} \sqrt{p_{\mathbf{x}}}|\mathbf{x}\rangle|\mathbf{x}\rangle$, where $p_{\mathbf{x}} \triangleq \Pr[(\mathbf{X}_i[\ell] : \ell \neq j) = \mathbf{x} \mid \mathbf{D} = \hat{\mathbf{d}}]$; register $\mathsf{IN}_i[j]$ is set to $|u_i\rangle$. Then, $P_1, \ldots, P_t$ run the protocol Π'. Note that the registers $\widetilde{\mathsf{IN}}_i[j], 1 \leq i \leq t$ do not exist in Π_h.

We need to verify that protocol Π_h has two properties. First, that it is a δ-error protocol for h. For this we note that in Π_h, at all times, the state of the registers that were present in the original protocol Π (that is all registers except $\widetilde{\mathsf{IN}}_i$) is identical to their state when Π is run with input $\mathsf{embed}(\mathbf{X}, j, (u_1, \ldots, u_t))$ conditioned on the event $\mathbf{D} = \hat{\mathbf{d}}$. Since $\mathbf{X}$ collapses F to h, we conclude that Π_h computes $h(u_1, \ldots, u_t)$ with probability at least $1 - \delta$.

Second, we need to verify that $\mathsf{IL}(\Pi_h \mid (X, D)) \leq \frac{2ck}{n}$. We expand the left hand side of (3) using definition (2) of $I(\mathbf{d})$ and show that each term in it is at least the corresponding term in $\mathsf{IL}(\Pi \mid (X, D))$. For example, consider the term $I((X^i : \mathcal{P}^i) \mid D = d)$, $1 \leq i \leq k$ in the definition of $\mathsf{IL}(\Pi_h \mid (X, D))$. Note that the state of $(X^i, \mathcal{P}^i)$ in Π_h on input X conditioned on $D = d$ is identical to the state of $(\mathbf{X}^i[j], \mathcal{P}^i)$ in Π' with registers $\widetilde{\mathsf{IN}}_\ell[j]$, ℓ ranging over all parties except $\mathcal{P}^i$ traced out, when Π' is run on input $\mathbf{X}$ conditioned on $\mathbf{D} = \mathsf{embed}(\hat{\mathbf{d}}, j, d)$. It follows from

the monotonicity of mutual information that $I((X^i : \mathcal{P}^i) \mid D = d) \leq I((\mathbf{X}^i[j] : \mathcal{P}^i) \mid \mathbf{D} = \mathsf{embed}(\hat{\mathbf{d}}, j, d))$. We can thus conclude that $\mathsf{IL}(\Pi_h \mid (X, D)) \leq \frac{2ck}{n}$. ■

Let D be a random variable taking values in $\{1, \ldots, t\}$, with $\Pr[D = d] \triangleq k(d)/k$. Let $\mathcal{X}_1 = \cdots = \mathcal{X}_t \triangleq \{0, 1\}$. Let X_i be a random variable taking values in $\mathcal{X}_i$ and $X \triangleq (X_1, \ldots, X_t)$. When $D = d$, $\Pr[X_d = 0] = \Pr[X_d = 1] = 1/2$ and $\Pr[X_i = 0] = 1, i \neq d$. It is clear that D partitions X. Note that X^n collapses DISJ to AND (here DISJ denotes the promise t-party set disjointness problem and AND denotes the AND function on t bits). We now show a lower bound for the conditional information loss of AND under (X, D).

Lemma 3 *Let (X, D) be as above. Let $0 \leq \epsilon \leq 1/2$. Then, $\mathsf{IL}_\epsilon^{t,k}(\text{AND} \mid (X, D)) \geq \frac{(1-2\epsilon)^2}{8k}$.*

Proof: Let $\theta > 0$. Let Π be a t-party k-round ϵ-error quantum protocol with prior entanglement for AND with $\eta \triangleq \mathsf{IL}(\Pi \mid (X, D)) \leq \mathsf{IL}_\epsilon^{t,k}(\text{AND} \mid (X, D)) + \theta$. Consider the situation in Π just after the ith round of communication. For any $\mathbf{x} \in \{0, 1\}^t$, let $|\phi_{\mathbf{x}}^i\rangle$ be the global state vector of the qubits of $P_1, \ldots, P_t$ at this point in time, when protocol Π is started with input $X = \mathbf{x}$. Define $s(i) \triangleq I((X^i : \hat{P}^i) \mid D = \mathcal{P}^i)$. Then, $s(i) = \frac{k}{k(i)} \cdot I((X^i : \hat{P}^i) \mid D)$. Hence, $\eta = \sum_{i=1}^k s(i)$. Let $\mathbf{e}_i \in \{0, 1\}^t$ denote the vector which has an 1 in the ith coordinate and 0 everywhere else. Let $\mathbf{0}, \mathbf{1} \in \{0, 1\}^t$ denote the all-zeroes and all-ones vectors respectively. To keep our notation concise, for state vectors $|\phi\rangle$ and $|\psi\rangle$ we write $\||\phi\rangle - |\psi\rangle\|_t$ instead of $\||\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|\|_t$. By Lemma 1, there is a 'correction' unitary transformation V^i acting on the qubits in the possession of $\mathcal{P}^i$ just after round i such that

$$\left\| V^i |\phi_{\mathbf{0}}^i\rangle - |\phi_{\mathbf{e}_{\mathcal{P}^i}}^i\rangle \right\|_t \leq \sqrt{8s(i)}. \tag{4}$$

For any $1 \leq j \leq t$, let W_j^i denote the 'correction' unitary transformation of party $\mathcal{P}_j$ in the last round at or before round i when $\mathcal{P}_j$ was active. Then, $W_{\mathcal{P}^i}^i = V^i$. For any $j \neq j'$ W_j^i and $W_{j'}^i$ act on disjoint sets of qubits. Without loss of generality, $\mathcal{P}^i = \mathcal{P}_1$ and $\mathcal{P}^{i+1} = \mathcal{P}_2$. Define $\delta_i \triangleq \left\| W_1^i W_3^i \cdots W_t^i |\phi_{\mathbf{e}_2}^i\rangle - |\phi_{\mathbf{1}}^i\rangle \right\|_t$. Let U^i denote the unitary transformation of protocol Π that $\mathcal{P}_1$ applies to the qubits in her possession just after round $i-1$ in order to prepare the messages of round i. Let i' denote the last round before i when $\mathcal{P}_2$ was active. Let $U^{i',i}$ denote the product of the unitary transformations applied by the parties in protocol Π after round i' is complete and till the end of round i. Then, $|\phi_{\mathbf{x}}^i\rangle = U^i |\phi_{\mathbf{x}}^{i-1}\rangle$ and $|\phi_{\mathbf{x}}^i\rangle = U^{i',i} |\phi_{\mathbf{x}}^{i'}\rangle$. For $j \neq 1$, U^i and W_j^i act on disjoint sets of qubits and $W_j^i = W_j^{i-1}$. Also, W_2^i and $U^{i',i}$ act on disjoint sets of qubits. Using the unitary invariance and triangle inequality of the trace distance, the fact that unitary transformations on disjoint sets of qubits commute, and (4),

$$\begin{aligned}
\delta_i &\triangleq \left\| W_1^i W_3^i \cdots W_t^i |\phi_{\mathbf{e}_2}^i\rangle - |\phi_{\mathbf{1}}^i\rangle \right\|_t \\
&\leq \left\| \begin{array}{l} W_1^i W_3^i \cdots W_t^i U^{i',i} |\phi_{\mathbf{e}_2}^{i'}\rangle \\ -W_1^i W_3^i \cdots W_t^i U^{i',i} W_2^i |\phi_{\mathbf{0}}^{i'}\rangle \end{array} \right\|_t + \\
&\quad \left\| \begin{array}{l} W_1^i W_3^i \cdots W_t^i U^{i',i} W_2^i |\phi_{\mathbf{0}}^{i'}\rangle \\ -U^i W_2^i \cdots W_t^i |\phi_{\mathbf{e}_1}^{i-1}\rangle \end{array} \right\|_t + \\
&\quad \left\| U^i W_2^i \cdots W_t^i |\phi_{\mathbf{e}_1}^{i-1}\rangle - U^i |\phi_{\mathbf{1}}^{i-1}\rangle \right\|_t \\
&= \left\| |\phi_{\mathbf{e}_2}^{i'}\rangle - W_2^i |\phi_{\mathbf{0}}^{i'}\rangle \right\|_t \\
&\quad + \left\| W_1^i U^{i',i} |\phi_{\mathbf{0}}^{i'}\rangle - U^i |\phi_{\mathbf{e}_1}^{i-1}\rangle \right\|_t \\
&\quad + \left\| W_2^i \cdots W_t^i |\phi_{\mathbf{e}_1}^{i-1}\rangle - |\phi_{\mathbf{1}}^{i-1}\rangle \right\|_t \\
&= \left\| |\phi_{\mathbf{e}_2}^{i'}\rangle - W_2^i |\phi_{\mathbf{0}}^{i'}\rangle \right\|_t + \left\| W_1^i |\phi_{\mathbf{0}}^i\rangle - |\phi_{\mathbf{e}_1}^i\rangle \right\|_t \\
&\quad + \left\| W_2^{i-1} \cdots W_t^{i-1} |\phi_{\mathbf{e}_1}^{i-1}\rangle - |\phi_{\mathbf{1}}^{i-1}\rangle \right\|_t \\
&\leq \sqrt{8s(i')} + \sqrt{8s(i)} + \delta_{i-1}.
\end{aligned}$$

It is easy to check that $\delta_0 = 0$. Hence, $\delta_k \leq \sum_{i=1}^k 2\sqrt{8s(i)}$. Using concavity of the square root function, we get that $\delta_k \leq \sqrt{32\eta k}$. Using Fact 2, the fact that a local unitary transformation does not affect the density matrix of the remote system and monotonicity of trace distance, we get that a correct k-round ϵ-error protocol for AND must have $\delta_k \geq 2 - 4\epsilon$. Hence, $\eta \geq \frac{(1-2\epsilon)^2}{8k}$ implying that $\mathsf{IL}_\epsilon^{t,k}(\text{AND} \mid (X, D)) \geq \frac{(1-2\epsilon)^2}{8k} - \theta$ for any $\theta > 0$. This completes the proof of the lemma. ■

Remark: In fact, there are bounded error two-party k-round quantum protocols for the AND of two bits with conditional information loss $O(\log k/k)$. Such protocols can be obtained from the protocols of [BCW98, HdW02, AA03] for set disjointness on a universe of size $O(k^2)$ by setting the first coordinate of Alice and Bob to the two input bits and setting the rest of the coordinates to 0. Another such protocol with one qubit messages can be obtained by adapting the 'reflections in a plane' visualisation (see e.g. [NC00]) of Grover's algorithm on a universe of size $O(k^2)$.

The following is now immediate from Lemma 2 and Lemma 3.

Theorem 1 *Any t-party k-round bounded error quantum protocol for the set disjointness problem needs to have communication cost at least $\Omega(n/k^2)$.*

Acknowledgements

Our original proof gave a lower bound of $\Omega(n/k^4)$ for set disjointness. We later improved it to $\Omega(n/k^2)$ using an

inequality implicitly contained in a paper by Lin [Lin91]. Hartmut Klauck independently pointed out to us that similar improvements can also be obtained using an inequality from [DHR78]. We thank him for sharing with us his insights and pointing reference [DHR78] to us. We also thank Andris Ambainis for pointing out reference [FvdG99], which contains an explicit form of Lin's inequality.

References

[AA03] S. Aaronson and A. Ambainis. Quantum search of spatial regions. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, 2003. To appear. Also quant-ph/0303041.

[AKN98] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998. Also quant-ph/9806029.

[AMS99] N. Alon, Y. Matias, and M. Szegedy. The space complexity of approximating the frequency moments. *Journal of Computer and System Sciences*, 58(1):137–147, 1999.

[BCW98] H. W. Buhrman, R. Cleve, and Avi Wigderson. Quantum vs classical communication and computation. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 63–68, 1998. Also quant-ph/9702040.

[BdW01] H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In *Proceedings of the 16th IEEE Conference on Computational Complexity*, pages 120–130, 2001.

[BFS86] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity theory. In *Proceedings of the 27th Annual IEEE Symposium on Foundations of Computer Science*, pages 337–347, 1986.

[BJKS02a] Z. Bar-Yossef, T. Jayram, R. Kumar, and D. Sivakumar. Information theory methods in communication complexity. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity*, pages 93–102, 2002.

[BJKS02b] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 209–218, 2002.

[CG88] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal of Computing*, 17(2):230–261, 1988.

[CKS03] A. Chakrabarti, S. Khot, and X. Sun. Near-optimal lower bounds on the multiparty communication complexity of set-disjointness. In *Proceedings of the 18th Annual IEEE Conference on Computational Complexity*, 2003.

[CSWY01] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 270–278, 2001.

[CvDNT98] R. Cleve, Wim van Dam, M. Nielsen, and A. Tapp. Quantum entanglement and the communication complexity of the inner product function. In *Proceedings of the 1st NASA International Conference on Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science, vol. 1509, pages 61–74. Springer-Verlag, 1998. Also quant-ph/9708019.

[DHR78] D. Dacunha-Castelle, H. Heyer, and B. Roynette. *Ecole d'Eté de Probabilités de Saint-Flour VII*. Lecture Notes in Mathematics, vol. 678. Springer-Verlag, 1978.

[FKS02] J. Feigenbaum, S. Kannan, and M. Strauss. An approximate l^1-difference algorithm for massive data streams. *SIAM Journal of Computing*, 32:131–151, 2002.

[FvdG99] C. A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999. Also quant-ph/9712042.

[GGI+02] A. Gilbert, S. Guha, P. Indyk, Y. Kotidis, S. Muthukrishnan, and M. Strauss. Fast small space algorithms for approximate histogram maintenance. In *Proceedings of the 34th Annual ACM Symposium Theory of Computing*, pages 389–398, 2002.

[GMMO00] S. Guha, N. Mishra, Rajeev Motwani, and L. O'Callaghan. Clustering data streams. In

Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science, pages 359–366, 2000.

[HdW02] P. Hoyer and R. de Wolf. Improved quantum communication complexity bounds for disjointness and equality. In *Symposium on Theoretical Aspects of Computer Science*, pages 299–310, 2002. Also quant-ph/0109068.

[Ind00] P. Indyk. Stable distributions, pseudorandom generators, embeddings, and data stream computations. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pages 189–197, 2000.

[JKS03] T.S. Jayram, R. Kumar, and D. Sivakumar. Two applications of information complexity. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 673–682, 2003.

[KKN95] M. Karchmer, E. Kushilevitz, and N. Nisan. Fractional covers and communication complexity. *SIAM Journal on Discrete Mathematics*, 8(1):76–92, 1995.

[Kla01] H. Klauck. Lower bounds for quantum communication complexity. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 288–297, 2001. Also at quant-ph/0106160.

[KN97] E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, 1997.

[KNTZ01] H. Klauck, A. Nayak, Amnon Ta-Shma, and D. Zuckerman. Interaction in quantum communication and the complexity of set disjointness. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 124–133, 2001.

[Kre95] I. Kremer. Quantum communication. Master's thesis, Hebrew University, 1995.

[KS92] Bala Kalyansundaram and G. Schnitger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992.

[Lin91] J Lin. Divergence measures based on Shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991.

[LS81] R.J. Lipton and R. Sedgewick. Lower bounds for VLSI. In *Proceedings of the 13th Annual ACM Symposium on Theory of Computing*, pages 300–307, 1981.

[MNSW98] P. Bro Miltersen, Noam Nisan, S. Safra, and A. 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 *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 369–376, 1999.

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Nis02] N. Nisan. The communication complexity of approximate set packing and covering. In *Proceedings of the 29th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, vol. 2380, pages 868–875, 2002.

[Raz92] A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[Raz02] A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya Math*, 6, 2002. In Russian. To appear. English version at quant-ph/0204025.

[Sen03] P. Sen. Lower bounds for predecessor searching in the cell probe model. In *Proceedings of the 18th Annual IEEE Conference on Computational Complexity*, 2003.

[SS02] M. Saks and X. Sun. Space lower bounds for distance approximation in the data stream model. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 360–369, 2002.

[Yao79] A. Yao. Some complexity questions related to distributed computing. In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.

[Yao93] A. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science*, pages 352–361, 1993.