Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 6 (2019)

Upper Bounds on Communication in terms of Approximate Rank

Anna Gál[*] Ridwan Syed[†]

Abstract

We show that any Boolean function with approximate rank r can be computed by bounded error quantum protocols without prior entanglement of complexity $O(\sqrt{r}\log r)$. In addition, we show that any Boolean function with approximate rank r and discrepancy δ can be computed by deterministic protocols of complexity $O(r)$, and private coin bounded error randomized protocols of complexity $O((\frac{1}{\delta})^2 + \log r)$. Our deterministic upper bound in terms of approximate rank is tight up to constant factors, and the dependence on discrepancy in our randomized upper bound is tight up to taking square-roots. Our results can be used to obtain lower bounds on approximate rank. We also obtain a strengthening of Newman's theorem with respect to approximate rank.

1 Introduction

The log-rank conjecture is one of the most intriguing open problems in communication complexity. Lovász and Saks [26] conjectured that the deterministic communication complexity $D(f)$ of a Boolean function f is upper bounded by $(\log rank(f))^k$ for some constant k, where $rank(f)$ denotes the rank over the reals of the communication matrix of the function f. It is a classical result in communication complexity by Mehlhorn and Schmidt [29] that $D(f) \geq \log rank(f)$. It is easy to see that $D(f) \leq rank(f) + 1$, but

[*]University of Texas at Austin, Email: panni@cs.utexas.edu, Part of this work was done while visiting the Simons Institute for the Theory of Computing in Berkeley.

[†]University of Texas at Austin, Email: ridwan@cs.utexas.edu

ISSN 1433-8092

until a few years ago no one was able to obtain upper bounds sublinear in $rank(f)$. The current best upper bound was obtained by Lovett [27] who proved that $D(f) \leq O(\sqrt{rank(f)} \log rank(f))$. Considering separations, a series of results (see [33]) showed that the deterministic communication complexity can be superlinear in the logarithm of $rank(f)$. Currently the largest known separation is nearly quadratic: [14] shows that there are functions with $D(f) \geq \tilde{\Omega}((\log rank(f))^2)$.

Krause [19] extended the log-rank lower bound method to the context of randomized communication complexity, by considering the approximate rank of communication matrices. Different definitions of approximate rank have been introduced in various contexts, see for example [3]. For $\alpha \geq 1$, the *α-approximate rank* of a matrix A with $+1, -1$ entries is the smallest possible rank of a matrix B that sign represents the matrix A (e.g. each entry of B has the same sign as the corresponding entry in A) with the additional condition that the absolute values of the entries are at least 1 and at most α. Krause [19] proved that the logarithm of α-approximate rank provides lower bounds on bounded error randomized communication complexity with private coins, allowing error at most $\epsilon = \frac{1}{2} - \frac{1}{2\alpha}$. The logarithm of approximate rank also lower bounds quantum communication [10]. We state the precise bounds in Section 2.

In light of these lower bounds, it is natural to consider the analogue of the log-rank conjecture for approximate rank in the context of randomized communication, and quantum communication. The "log-approximate-rank" conjecture and its quantum version were stated in the survey by Lee and Shraibman [21]. The conjecture states that $R_\epsilon(f) \leq (\log \mathrm{rk}^\alpha(f))^k$ for $\epsilon = \frac{1}{2} - \frac{1}{2\alpha}$ and some constant k, where R_ϵ denotes private coin randomized communication with error ϵ and rk^α denotes α-approximate rank.

If we relax the definition of approximate rank and allow arbitrary real entries in the matrix that sign represents the communication matrix, then we obtain the definition of "sign rank" (sometimes referred to as "dimension complexity"), which is a well studied measure. The appropriate communication model to consider is "unbounded error communication complexity", where it is only required that the error probability is less than 1/2. Paturi and Simon [34] proved that the unbounded error communication complexity essentially equals to the logarithm of sign rank of the function, thus in this model the corresponding version of the log-rank conjecture holds.

In the case of randomized and quantum communication complexity, with error bounded by a constant $\epsilon < 1/2$, the conjecture has been recently refuted

[11, 5, 38]. Until recently, the largest known separation between the logarithm of approximate rank and quantum communication has been nearly quadratic [4]. A quadratic separation between the logarithm of approximate rank and randomized communication is witnessed by the disjointness function. The n-bit disjointness function can be computed by $O(\sqrt{n})$-qubit quantum protocols [1] and thus it has approximate rank $2^{O(\sqrt{n})}$ by the lower bound of [10] on quantum communication. On the other hand it requires $\Omega(n)$ randomized communication [17, 35]. A breakthrough result of Chattopadhyay et al. [11] gave an example of a Boolean function with approximate rank r that requires $\Omega(r^{1/4})$ randomized communication. [5, 38] showed that the same function requires $\Omega(r^{1/12})$ quantum communication.

1.1 Our Results

In this paper we consider upper bounds on communication in terms of approximate rank. Approximate rank can be significantly smaller than rank. For example the equality function gives an exponential separation: the approximate rank of the n-bit equality function is $\Theta(n)$ while its rank is 2^n. Thus, upper bounds on communication complexity in terms of approximate rank can potentially give much sharper upper bounds on communication complexity. Our randomized upper bounds also involve the inverse of discrepancy, which in turn can be arbitrarily small compared to approximate rank: the discrepancy of the equality function is constant.

Our main results are the following upper bounds: We show that any Boolean function with approximate rank r and discrepancy δ can be computed by

- deterministic protocols of complexity $O(r)$,
- private coin bounded error randomized protocols of $O((\frac{1}{\delta})^2 + \log r)$ complexity and
- bounded error quantum protocols without prior entanglement of complexity $O(\sqrt{r} \log r)$.

The example of the equality function shows that our deterministic upper bound in terms of approximate rank is tight up to constant factors, and that the additive $\log r$ term is necessary in our randomized upper bound. In addition, the disjointness function shows that the dependence on discrepancy in our randomized bound is tight up to taking square-roots.

While in the deterministic case the upper bound $D(f) \leq rank(f) + 1$ is immediate from bounding the number of different rows of the matrix, as far as we know, a linear upper bound on communication in terms of approximate rank alone has not been established before, considering deterministic or private coin bounded-error randomized communication. We show that any Boolean function with α-approximate rank r can be computed by deterministic protocols of complexity $O(r)$, where the constant in the big-Oh notation depends on α. This bound in turn can also be used to obtain lower bounds on approximate rank (see Section 3.1).

Lovett [27] asked if the techniques developed to prove his result can be generalized to obtain upper bounds on randomized and quantum communication complexity in terms of approximate rank. Lovett's result [27] implies that deterministic communication (and hence randomized communication) is upper bounded by square-root of approximate rank with an additional $(\log rank(f))^2$ factor. This follows since Lovett's proof gives that $D(f) \leq O(\frac{1}{disc(f)}(\log rank(f))^2)$, where $disc(f)$ is the discrepancy of f, and since $\frac{1}{disc(f)}$ is bounded above by the square-root of approximate rank [25]. However, when approximate rank is much smaller than $rank(f)$, this bound may be superlinear in approximate rank.

Results of Linial and Shraibman (Claim 2. in [25]), and Klauck [18] imply public coin bounded error randomized protocols with communication $O((\frac{1}{disc(f)})^2)$ and therefore with communication linear in approximate rank. It is well known by a theorem of Newman [32] that public coin protocols can be converted to private coin protocols (with slightly larger error) at the cost of additional $O(\log n + \log(1/\rho))$ bits where ρ is the (additive) increase in error. However, applying Newman's theorem as stated, does not give our claimed bound when the approximate rank is very small. We obtain our result on private coin protocols by giving a strengthening of Newman's theorem with respect to approximate rank.

Considering quantum protocols (without prior entanglement) we are able to match Lovett's bound in terms of approximate rank. We show that any Boolean function with α-approximate rank r can be computed by quantum protocols of complexity $O(\alpha^2\sqrt{r}\log r)$. To obtain these bounds, we first show that any function with α-approximate rank r can be computed by $O(\log r)$ communication and error at most $\frac{1}{2} - \frac{1}{2\alpha\sqrt{r}}$ by private coin randomized or quantum protocols. Previously, private coin protocols with $O(\log r)$ communication but with error $\frac{1}{2} - \frac{1}{2\alpha r}$ were given in [12]. Even with our improved error bound, amplifying the correctness of this protocol classically would only give bounded error private coin protocols with $O(r \log r)$ communication. Using the amplitude amplification technique of [30] allows us to obtain bounded error quantum protocols with $O(\alpha^2 \sqrt{r} \log r)$ communication.

2 Preliminaries

Let $f : X \times Y \to \{-1, 1\}$ be a Boolean function. We write $D(f)$, $R_\epsilon(f)$, $R_\epsilon^{pub}(f)$, $U(f)$ to respectively denote the deterministic communication complexity, ϵ-error randomized communication complexity with private coins, ϵ-error randomized communication complexity with public coins, and unbounded error randomized communication complexity with private coins. For background in classical communication complexity we refer to [20]. For the quantum analogues of these measures we write $Q(f), Q_\epsilon(f), Q_\epsilon^*(f)$ to respectively denote the exact quantum communication complexity, ϵ-error quantum communication complexity without prior entanglement, and ϵ-error quantum communication complexity with prior entanglement. We postpone a brief review of the quantum communication model until Section 5.

We will often identify f with its communication matrix M with entries $M[x, y] = f(x, y)$. Our primary complexity measure of interest is the approximate rank.

Definition 1. *[19](see also [21]) Let $f : X \times Y \to \{-1, 1\}$ and let M be the communication matrix of f. Fix some real $\alpha \geq 1$. The α-approximate rank of f is defined as*

$$rk^\alpha(f) := \min_{A \,:\, 1 \leq A_{i,j} \cdot M_{i,j} \leq \alpha} rk(A)$$

where rk is the usual rank over $\mathbb{R}$, and the entries of A are reals. We say such a matrix A is an approximating matrix for f. Note that when $\alpha = 1$, this measure coincides with the usual rank of f. When we do not bound the entries of A, this measure coincides with the sign rank of f, denoted $rk^\infty(f)$.

We say that a matrix B sign represents a communication matrix M if $B_{i,j} M_{i,j} > 0$ for all i, j. Note that an approximating matrix for f sign represents the communication matrix of f.

Sometimes approximate rank of Boolean functions is defined with respect to 0/1 matrices and parameter $0 \leq \epsilon < 1/2$. Taking $\epsilon = \frac{1}{2} - \frac{1}{2\alpha}$ the two definitions are equivalent (up to additive +1 or -1).

For a matrix A, we call a decomposition of the form $A = UV$ a d-dimensional factorization if the rows (resp. columns) of U (resp. V) are in $\mathbb{R}^d$. Recall that the rank of a matrix is the minimum d for which such a factorization is possible.

We will also be interested in bounds on the lengths of vectors in such decompositions, which is captured by factorization norms. Approximate factorization norm (γ_2 norm) was introduced by Linial and Shraibman [25], who showed that the logarithm of α-approximate γ_2 norm is a lower bound for quantum communication with entanglement and error $\epsilon = \frac{1}{2} - \frac{1}{2\alpha}$.

Definition 2. *[25] Let $f : X \times Y \to \{-1, 1\}$ be a Boolean function, and let M be its communication matrix. Fix some real $\alpha \geq 1$. The α-approximate factorization norm of f is defined as*

$$\gamma_2^{\alpha}(f) = \min_{A=UV \ : \ 1 \leq A_{i,j} \cdot M_{i,j} \leq \alpha} \ell(U) \cdot \ell(V^T)$$

where $\ell(B)$ denotes the maximum ℓ_2 norm of a row of B. When $\alpha = 1$, we omit the superscript and simply write $\gamma_2(f)$. When we do not bound the entries of A, we write $\gamma_2^{\infty}(f)$.

The approximate γ_2 norm can be significantly smaller than the approximate rank. For example the γ_2 norm of the n-bit equality function is constant, while its approximate rank is $\Omega(n)$ [2, 22].

Our results exploit the following relationship shown by [22] between γ_2^{α} and rk^{α}. See also [23, 37] for more on relating factorization norm and rank.

Theorem 1. *For any $f : X \times Y \to \{-1, 1\}$ and $\alpha > 1$,*

$$\gamma_2^{\alpha}(f) \leq \alpha \sqrt{rk^{\alpha}(f)}$$

Moreover this bound can be witnessed by an approximating matrix $A = UV$ with a factorization of dimension $rk^{\alpha}(f)$ such that $\ell(U) \leq \sqrt{rk^{\alpha}(f)}$ and $\ell(V) \leq \alpha$.

For this factorization, we will find it convenient to explicitly refer to the rows (resp. columns) of U (resp. V), and we will additionally enforce uniformity in the respective ℓ_2 norms by slightly increasing dimension.

Lemma 1. *Let $f : X \times Y \to \{-1, 1\}$ be a Boolean function with $rk^{\alpha}(f) = r$. There exist factorization vectors $\{u_x\}_{x \in X}$ and $\{v_y\}_{y \in Y}$ in $\mathbb{R}^{r+2}$ such that for all x, y, $\|u_x\| = \sqrt{r}$, $\|v_y\| = \alpha$, and $1 \leq (u_x \cdot v_y) f(x, y) \leq \alpha$.*

Proof. Let $A = UV$ be the factorization guaranteed by Theorem 1. Write the row of U corresponding to $x \in X$ as u'_x. Let $s_x = r - \|u'_x\|_2^2$. Write the column of V corresponding to $y \in Y$ as v'_y. Let $s_y = \alpha^2 - \|v'_y\|_2^2$. Finally define $u_x = u'_x \oplus (s_x, 0)$ and $v_y = v'_y \oplus (0, s_y)$. It is straightforward to check that these vectors satisfy the claims of the lemma. □

We record the lower bounds alluded to in the introduction. For $\frac{1}{2} > \epsilon > 0$ let $\alpha = \frac{1}{1-2\epsilon}$. We use log to denote $\log_2$ unless otherwise indicated.

- $D(f) \geq \log \mathrm{rk}(f) \geq \log \gamma_2(f)$ [29, 23]
- $R_\epsilon(f) \geq \log \mathrm{rk}^\alpha(f) \geq 2 \log \gamma_2^\alpha(f) - 2 \log \alpha$ [19, 25]
- $U(f) \geq \log \mathrm{rk}^\infty(f)$ [34]
- $Q(f) \geq \log \mathrm{rk}(f)/2$ [9, 10]
- $Q_\epsilon(f) \geq \log \mathrm{rk}^\alpha(f)/2$ [10]
- $Q_\epsilon^*(f) \geq \log \gamma_2^\alpha(f) - \log \alpha - 2$ [25]

Another important measure we consider is discrepancy.

Definition 3. *For a Boolean function $f : X \times Y \to \{-1, 1\}$ its discrepancy is defined as*

$$disc(f) = \min_\mu \max_R |\sum_{(x,y)\in R} f(x,y)\mu(x,y)|$$

where μ is an arbitrary distribution over $X \times Y$ and $R = A \times B$ with $A \subseteq X$, $B \subseteq Y$ is an arbitrary rectangle.

It is known that $\frac{1}{disc(f)}$ is equivalent to $\gamma_2^\infty(f)$ up to constant factors [24]. Thus, by Theorem 1 for constant $\alpha \geq 1$ and any Boolean function f we have

$$\Theta(\frac{1}{disc(f)}) = \gamma_2^\infty(f) \leq \gamma_2^\alpha(f) \leq \alpha\sqrt{\mathrm{rk}^\alpha(f)} = O(\sqrt{\mathrm{rk}(f)})\,. \qquad (1)$$

For more on communication lower bounds and the relationships between these and other measures we refer to an excellent survey by Lee and Shraibman [21].

For public coin protocols, the following upper bounds have been established in terms of discrepancy.

Lemma 2. *[25] Let $f : X \times Y \to \{1, -1\}$ be a Boolean function. Then there is an $O(1)$-bit public coin randomized protocol for f with error at most $\frac{1}{2} - \frac{1}{2K\gamma_2^\infty(f)}$, where $1.5 \leq K \leq 1.8$ is the Grothendieck constant.*

Since $\frac{1}{disc(f)}$ is equivalent to $\gamma_2^\infty(f)$ up to constant factors [24], this also gives a bound with respect to discrepancy. A more direct argument in terms of discrepancy is given by Klauck [18].

Lemma 3. *[18] Let $f : X \times Y \to \{1, -1\}$ be a Boolean function. Then there is an $O(1)$-bit public coin randomized protocol for f with error at most $\frac{1}{2} - \frac{disc(f)}{2}$.*

The following upper bound on public coin randomized protocols follows by standard amplification.

Theorem 2. *[25, 18] Let $f : X \times Y \to \{1, -1\}$ be a Boolean function. Then $R_{1/3}^{pub}(f) = O((\frac{1}{disc(f)})^2)$.*

The following private coin protocol with $O(\log \mathrm{rk}^\alpha(f))$ communication was observed in [12]. This protocol is similar to the proof of Paturi and Simon [34] in the unbounded error model.

Lemma 4. *[12] Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and $\alpha > 1$. Then there is a $\log(4rk^\alpha(f))$-bit private coin randomized protocol with error at most $\frac{1}{2} - \frac{1}{2\alpha rk^\alpha(f)}$,*

For functions f with $\mathrm{rk}^\alpha(f) = r$, amplifying this protocol gives private coin bounded error protocols with $O(\alpha^2 r^2 \log r)$ communication. We improve the error bounds of the protocol of Lemma 4 and we use the existence of this protocol as a starting point for our results.

3 Deterministic Communication

In this section we give upper bounds on the number of different rows and columns of Boolean matrices with given approximate rank. These estimates directly yield upper bounds on deterministic communication in terms of approximate rank.

We need the following theorem of Krause [19].

Theorem 3. *[19] Let $f : X \times Y \to \{1, -1\}$ be a Boolean function, and let M be the communication matrix of f. If there is a private coin randomized protocol for f with c-bit communication and error at most $\frac{1}{2} - \frac{1}{s}$, then there is a $2^n \times 2^n$ matrix B with nonzero integer entries that sign represents M such that the absolute values of the entries of B are at most $t = 8s2^c$, and B has rank at most 2^c.*

Note that the lower bound of Krause stated in the Introduction and Section 2 is a simplified version of this theorem, without requiring that the approximating matrix has integer elements.

Theorem 3 allows us to estimate the number of different rows and columns in the communication matrix of a function with given approximate rank.

Lemma 5. *Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and $\alpha \geq 1$. Let M be the communication matrix of the function f. If $rk^{\alpha}(f) = r$, then M has at most $(2t)^{4r}$ distinct rows and columns, where $t = 64\alpha r^2$.*

Proof. Applying Theorem 3 to the protocol from Lemma 4 we know that there is a $2^n \times 2^n$ matrix B with nonzero integer entries that sign represents M such that the absolute values of the entries of B are at most $t = 64\alpha r^2$, and B has rank at most $4r$. Thus the number of different rows and columns of B is at most $(2t)^{4r}$. But since B sign represents M, the number of different rows (and columns) of M cannot be larger than the number of different rows (and columns) of B. □

Shachar Lovett [28] pointed out to us that our bound can be improved as follows.

Lemma 6. *[28] Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and $\alpha \geq 1$. Let M be the communication matrix of the function f. If $rk^{\alpha}(f) = r$, then M has at most $(\alpha + 2)^r$ distinct rows and columns.*

Proof. Let $A \in R^{|X| \times |Y|}$ be an approximating matrix for f. Let V denote the linear span of the rows of A. Consider a maximal set of pairwise different rows of M, and denote the corresponding rows of A by $v_1, \ldots, v_N$.

We will use the notation $W = V \cap [-\alpha, \alpha]^{|Y|}$. By definition, $v_i \in W$ for each $i \in [N]$. Moreover, for $i_1 \neq i_2$ there is an index $1 \leq j \leq |Y|$ such that $|v_{i_1}(j) - v_{i_2}(j)| \geq 2$, where $v_i(j)$ denotes the j-th coordinate of the vector v_i.

We also use the notation $\beta W = \{\beta w | w \in W\}$ for constant $\beta > 0$, and we will consider $v + \beta W = \{v + u | u \in \beta W\}$.

Let us fix $\beta = \frac{1}{\alpha+1}$. With this notation, we have that the sets $v_i + \beta W$ are pairwise disjoint for $i \in [N]$. Thus, $\sum_{i\in[N]} \mathrm{Vol}(v_i + \beta W) = N\mathrm{Vol}(\beta W) \leq \mathrm{Vol}((1+\beta)W)$. This implies that $N \leq (\frac{1+\beta}{\beta})^r = (\alpha+2)^r$. $\square$

The above bounds on the number of different rows and columns imply the following upper bound on deterministic communication in terms of approximate rank.

Theorem 4. *Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and $\alpha \geq 1$. Then $D(f) \leq 1 + c_\alpha rk^\alpha(f)$, where $c_\alpha = \log(\alpha+2)$.*

Note that this bound is tight up to constant factors, as demonstrated by the equality function.

3.1 Lower Bounds on Approximate Rank

Theorem 4 can be used to derive lower bounds on approximate rank. First we note that it implies that for constant α the separation between approximate rank and rank is at most exponential. This also holds with respect to the nonnegative rank $\mathrm{rk}^+(f)$ of the corresponding 0/1 matrix.

Corollary 1. *Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and let $\alpha > 1$ be a constant. Then $rk^\alpha(f) \geq \Omega(\log rk^+(f)) \geq \Omega(\log rk(f))$.*

Next we note that by Theorem 4 any lower bound on the deterministic communication complexity of a function yields lower bounds on its approximate rank. Previous lower bounds on approximate rank are usually based on measures like discrepancy, factorization norm and trace norm. Our method will not give larger than n lower bounds for communication problems involving n-bit inputs for the players, but it can give interesting lower bounds for functions where the previously used measures are not very large.

For the n-bit Greater Than function GT_n Braverman and Weinstein [7] proved that $\frac{1}{disc(GT_n)} \geq \sqrt{n}$. This implies using (1) that for constant α the approximate rank of GT_n is $\Omega(n)$. Theorem 4 provides another proof of the $\Omega(n)$ lower bound on the approximate rank of GT_n, with improved dependence on α.

Corollary 2. *For $\alpha > 1$, $rk^\alpha(GT_n) \geq \frac{n}{\log(\alpha+2)}$.*

We also obtain a new proof of the $\Omega(n)$ lower bound of Alon [2] on the approximate rank of the n-bit Equality function EQ_n.

Corollary 3. *For $\alpha > 1$, $rk^\alpha(EQ_n) \geq \frac{n}{\log(\alpha+2)}$.*

4 Randomized Communication

Using Lemma 5 we obtain the following strengthening of Newman's theorem [32] with respect to approximate rank.

Theorem 5. *Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and $\alpha \geq 1$. For every $\rho > 0$ and every $\epsilon > 0$, $R_{\epsilon+\rho}(f) \leq R^{pub}_{\epsilon}(f) + O(\log rk^{\alpha}(f) + \log(\alpha + 2) + \log \rho^{-1})$.*

Proof. Let M be the communication matrix of the function. If the number of different rows of the matrix M is at most N_1 and the number of different columns of M is at most N_2 then the players can determine the value of f by running a protocol for a function $f' : X' \times Y' \to \{1, -1\}$ with $|X'| = N_1$ and $|Y'| = N_2$. Newman's theorem [32] gives that $R_{\epsilon+\rho}(f) \leq R^{pub}_{\epsilon}(f) + O(\log \log(N_1 N_2) + \log \rho^{-1})$, where N_1 is the number of different rows of M and N_2 is the number of different columns of M. The statement follows by Lemma 5. □

Theorem 5 allows us to simulate public coin protocols by private coin protocols efficiently with respect to approximate rank. We obtain the following bounds on private coin protocols.

Theorem 6. *Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and $\alpha \geq 1$. There is a private coin protocol computing f with $O(\log rk^{\alpha}(f) + \log(\alpha + 2))$ communication and error at most $\frac{1}{2} - \frac{disc(f)}{4}$.*

Proof. Follows by applying Theorem 5 with $\rho = \frac{disc(f)}{4}$ to the protocol given by Lemma 3, and using that $\frac{1}{disc(f)} = O(\alpha\sqrt{\mathrm{rk}^{\alpha}(f)})$ by (1). □

Theorem 7. *Let $f : X \times Y \to \{1, -1\}$ be a Boolean function and $\alpha \geq 1$. Then $R_{1/3}(f) = O((\frac{1}{disc(f)})^2 + \log rk^{\alpha}(f) + \log(\alpha + 2))$.*

Proof. Follows by applying Theorem 5 to Theorem 2. □

The equality function shows that the additive $\log \mathrm{rk}^{\alpha}(f)$ term is necessary in Theorem 7, since $\frac{1}{disc(EQ_n)} = O(\gamma_2(EQ_n)) = O(1)$. The disjointness function shows that the dependence on $\frac{1}{disc(f)}$ in Theorem 7 is tight up to taking square-roots: it is known that $\frac{1}{disc(DISJ_n)} = O(n)$ [20], and $\log \mathrm{rk}^{\alpha}(DISJ_n) = O(\sqrt{n})$ by combining the existence of $O(\sqrt{n})$-qubit quantum protocols [1] for $DISJ_n$ and the lower bound of [10] on quantum protocols as noted in the introduction. On the other hand $DISJ_n$ requires $\Omega(n)$ randomized communication [17, 35].

5 Quantum Communication

5.1 Quantum Communication Model

We assume basic familiarity with quantum information, and refer to [31] for more background. The state space of a quantum communication protocol is comprised of three registers : Alice's private register, Bob's private register, and a shared register. Each of these registers may be of some arbitrary fixed size which may depend on the function (but not the inputs) being computed. We assume that Alice and Bob are given their inputs x, y encoded as computational basis states $|x\rangle, |y\rangle$, and that the initial state of the protocol is

$$|\text{init}\rangle = |x\rangle \left|\vec{0}\right\rangle |y\rangle$$

Alice and Bob then alternate sending each other messages across the channel. More precisely, when it is Alice's ith turn to speak she applies some unitary U^{A_i} to her register and the shared register, and when it is Bob's jth turn to speak he applies some unitary U^{B_j} to his register and the shared register. At the end of the protocol, one of the players measures the qubits in the channel and as a function of the result outputs a result for the protocol. Recall that we write $Q(f), Q_\epsilon(f)$ to respectively denote the exact quantum communication complexity and ϵ-error quantum communication complexity and ϵ-error quantum communication.

5.2 Root Approximate Rank Upper Bound

Let $\text{rk}^\alpha(f) = r$ and assume as well that $\alpha > 1$ is a constant. Our main protocol combines two standard techniques: quantum fingerprinting[1] and amplitude amplification. The first, is for Alice and Bob to associate their (potentially long) inputs x and y with significantly shorter quantum states or *fingerprints* [8]. In particular, we will have Alice and Bob associate their inputs with quantum states which encode the factorization vectors $u_x, v_y \in \mathbb{R}^{O(r)}$ given by Lemma 1. In particular, such vectors (up to normalization) can be encoded with $O(\log r)$ qubits. Alice and Bob can then perform a distributed variant of the Hadamard test on their encoded states, which will

[1] Our protocol does not use fingerprinting in the usual sense since we will really be encoding the factorization vectors corresponding to Alice and Bob's inputs rather than their actual inputs.

allow them to estimate the inner product of their respective fingerprints, and in turn compute $f(x,y)$ with bias $1/O(\sqrt{r})$ over random guessing. To improve the bias to a constant classically requires $\Omega(r)$ repetitions.

To improve the bias more efficiently we apply the second technique : amplitude amplification. Generalizing the ideas involved in Grover's search algorithm [15], the technique of amplitude amplification [30] has been applied to quantum search and several problems in communication complexity [16, 9] to achieve polynomial speedups over classical algorithms. We record the technique in the following lemma.

Lemma 7. *Let $\mathcal{A}$ be a quantum algorithm which makes no measurements, and let G be some collection of 'good' basis states. Suppose that on initial state $|0\rangle$ the algorithm produces the state*

$$\mathcal{A}|0\rangle = \sin(\theta)|g\rangle + \cos(\theta)|b\rangle$$

where $|g\rangle$ is contained in a subspace W_G spanned by states in G, $|b\rangle$ is contained in the orthogonal complement of W_G, and $\theta \in [0, \pi/2]$. Let $\mathcal{B} = -\mathcal{A}S_0\mathcal{A}^{-1}S_G$, where S_G negates the amplitudes of the states in W_G, and S_0 negates the amplitude of $|0\rangle$. Repeatedly applying $\mathcal{B}$ to $\mathcal{A}|0\rangle$ k times produces the state

$$\mathcal{B}^k\mathcal{A}|0\rangle = \sin((2k+1)\theta)|g\rangle + \cos((2k+1)\theta)|b\rangle$$

We can view the pre-measurement state $\mathcal{A}|\text{init}\rangle$ of the protocol as a unit vector in a two dimensional space spanned by $|\text{g}\rangle$ (corresponding to the protocol outputting 1) and $|\text{b}\rangle$ (corresponding to the protocol outputting -1). As a two dimensional vector in this space, the vector $\mathcal{A}|\text{init}\rangle$ has angle $\theta = \pi/4+\delta$ with $|\text{b}\rangle$ if $f(x,y) = 1$, and angle $\theta = \pi/4-\delta$ with $|\text{b}\rangle$ if $f(x,y) = -1$, where $\delta > 0$ is some small value. The operator $\mathcal{B}$ above essentially implements a rotation of 2θ towards $|\text{g}\rangle$. Thus applying this operator 4 times, results in the vector having angle $\theta = \pi/4 + 8\delta$ with $|\text{b}\rangle$ if $f(x,y) = 1$, and angle $\theta = \pi/4 - 8\delta$ with $|\text{b}\rangle$ if $f(x,y) = -1$. This will correspond to increasing the bias by a constant factor! As we will show, Alice and Bob can apply $\mathcal{B}$ in blocks of 4 repetitions so that $O(\sqrt{r})$ rounds of application suffice to amplify the bias to a constant.

We begin with the fingerprinting protocol.

Lemma 8. *Fix $\alpha > 1$, and let $\epsilon = 1/2 - 1/(2\alpha\sqrt{r})$. If $rk^{\alpha}(f) = r$, then $Q_\epsilon(f) = O(\log r)$.*

Proof. Note that the statement follows from Theorem 6 with slightly different error bound. Here we give a direct proof by presenting an explicit protocol, which will be convenient for us to use in the proof of the next theorem.

The protocol we give is simply a distributed version of what is sometimes called the *Hadamard test.* Let x and y be Alice and Bob's respective inputs. By Lemma 1, there are vectors $u_x, v_y \in \mathbb{R}^{r+2}$ satisfying

$$\frac{1}{\alpha\sqrt{r}} \leq f(x,y)\frac{u_x \cdot v_y}{\|u_x\|_2\|v_y\|_2} \leq \frac{1}{\sqrt{r}} \tag{2}$$

The vectors' normalizations $u_x/\|u_x\|_2$ and $v_y/\|v_y\|_2$ can be encoded as $d = O(\log(r+2))$ qubit quantum states $|\phi_x\rangle, |\psi_y\rangle$ in a straightforward manner, so that the coordinates of the normalized vectors are precisely the amplitudes of the states in the computational basis. Alice and Bob's protocol is as follows:

1. Alice prepares the $d+1$ qubit state $\frac{1}{\sqrt{2}}|0\rangle|\phi_x\rangle + \frac{1}{\sqrt{2}}|0\rangle|0\rangle$ and sends the state to Bob.

2. Bob unitarily transforms the received state to $\frac{1}{\sqrt{2}}|0\rangle|\phi_x\rangle + \frac{1}{\sqrt{2}}|1\rangle|\psi_y\rangle$ and sends the state to Alice.

3. Alice performs a Hadamard transformation on the first qubit, and measures the first qubit. If she measures 0, Alice outputs 1, and otherwise she outputs -1.

We note that the protocol can be implemented so that it acts as the identity on each of Alice's and Bob's private registers. Clearly the total communication is $O(\log r)$. By a straightforward calculation, the probability that Alice measures 0 in the first qubit is

$$\frac{1}{2}(1 + \langle\phi_x|\psi_y|\phi_x|\psi_y\rangle)$$

The upper bound on the error probability of the protocol follows immediately from (2). □

We are now ready to prove our main theorem.

Theorem 8. *Fix $\alpha > 1$. If $rk^{\alpha}(f) = r$, then $Q_{1/3}(f) = O(\alpha^2\sqrt{r}\log r)$.*

Proof. Let $\mathcal{A}$ be the pre-measurement steps of the fingerprinting protocol of Lemma 8. Let the set of good states G be the set of basis states for which the first qubit of Alice and Bob's shared qubits is 0, and let W_G be the subspace spanned by these basis states. As in Lemma 7 let S_0 be an operator which negates the amplitude on the initial state[2] $|\text{init}\rangle = |x\rangle \left|\vec{0}\right\rangle |y\rangle$ of the protocol and let S_G be an operator which negates the amplitude on states in W_G. Finally let j, k be integers whose values we set later. Alice and Bob's protocol is as follows:

1. Alice and Bob apply $\mathcal{A}$ to the initial state $|\text{init}\rangle$.
2. Alice applies S_G to the shared qubits, negating the amplitudes on states in W_G.
3. Alice and Bob apply the inverse $\mathcal{A}^{-1}$ of the basic protocol.
4. Alice applies S_0 to the shared qubits, negating the amplitude on the initial state $|\text{init}\rangle$.
5. Alice and Bob apply $\mathcal{A}$, and then negate all amplitudes.
6. Alice and Bob repeat steps 2-5 an additional $4j - 1$ times so that the final state of the protocol is $\mathcal{B}^{4j}\mathcal{A}\,|\text{init}\rangle$, where $\mathcal{B} = -\mathcal{A}S_0\mathcal{A}S_G$.
7. Alice measures the first qubit. If she measures 0, Alice outputs 1, and otherwise she outputs -1.
8. Alice and Bob repeat steps 1-7 k independent times and output the majority result.

Clearly the communication cost of each of steps 1-5 is $O(\log r)$. Thus the total communication cost of the protocol is $O(jk \log r)$. It remains to set the parameters j, k.

By (2) the probability of measuring a state in W_G after step 1 is $1/2 + f(x, y) \cdot \delta$, where $\delta \in [1/(2\alpha\sqrt{r}), 1/(2\sqrt{r})]$. Thus we can write the state of the protocol after step 1 as

$$\mathcal{A}\,|\text{init}\rangle = \sin(\pi/4 + \theta)\,|\text{g}\rangle + \cos(\pi/4 + \theta)\,|\text{b}\rangle$$

[2] It suffices for S_0 to negate the amplitudes on states where all channel qubits are 0.

It follows from basic trigonometric identities that $\sin(2\theta) = 2f(x,y)\cdot\delta$. For all $z \in (0,1]$, $z < \sin^{-1}(z) \leq z\pi/2$. Thus, $1/(2\alpha\sqrt{r}) < f(x,y)\cdot\theta < \pi/(4\sqrt{r})$.

By Lemma 7 the state of the protocol after step 6 can be written as

$$\mathcal{B}^{4j}\mathcal{A}\left|\mathrm{init}\right\rangle = \sin(\pi/4 + (1+4j)\theta)\left|\mathrm{g}\right\rangle + \cos(\pi/4 + (1+4j)\theta)\left|\mathrm{b}\right\rangle \tag{3}$$

Take j to be the largest integer such that $(1+4j) < \sqrt{r}$. Given our choice of j, we can re-write 3 as

$$\mathcal{B}^{4j}\mathcal{A}\left|\mathrm{init}\right\rangle = \sin(\theta')\left|\mathrm{g}\right\rangle + \cos(\theta')\left|\mathrm{b}\right\rangle \tag{4}$$

where $\pi/4 + 1/(c\alpha) \leq \theta' \leq \pi/2$ for some constant c. It follows that the probability that Alice outputs $f(x,y)$ is at least $1/2 + 1/(c'\alpha)$ for some constant c'. By a standard argument we can take $k = O(\alpha^2)$, so that the entire protocol has cost $O(\alpha^2\sqrt{r}\log r)$ and the final result is correct with probability at least 2/3. □

When $\alpha = 1$ (that is when r is the rank of f) we can modify the analysis here to get an exact protocol. In particular, if u_x, v_y arise from a rank r factorization, then (2) can be written as the equality $(u_x \cdot v_y)/(\|u_x\|_2\|v_y\|_2) = f(x,y)/\sqrt{r}$. We can adjust the amount by which we increase the norms of u_x, v_y in the Lemma 1, so that in (4) for some integer $j = O(\sqrt{r})$ we have $(1+4j)\theta = f(x,y)\pi/4$. This gives an alternative and more direct proof of the following result which is a corollary of Lovett's deterministic upper bound in terms of rank.

Theorem 9. *If rk$(f) = r$, then $Q(f) = O(\sqrt{r}\log r)$.*

Acknowledgements

We thank Shachar Lovett for allowing us to include his argument improving our bound in Lemma 6. We would also like to thank the anonymous referees of a previous version of this paper for helpful comments. The second author thanks Patrick Rall and Daniel Liang for helpful discussions.

References

[1] S. Aaronson and A. Ambainis. Quantum search of spatial regions. Theory of Computing, (1) pp. 47–79, 2005.

[2] N. Alon. Perturbed identity matrices have high rank: proof and applications. Combinatorics, Probability, and Computing, (18) pp. 3–15, 2009.

[3] N. Alon, T. Lee, A. Shraibman, S. Vempala. The approximate rank of a matrix and its algorithmic applications. In Proceedings of STOC 2013, pp. 675-684.

[4] Anurag Anshu, Shalev Ben-David, Ankit Garg, Rahul Jain, Robin Kothari, Troy Lee. Separating Quantum Communication and Approximate Rank. Computational Complexity Conference 2017: 24:1–24:33.

[5] Anurag Anshu, Naresh Goud Boddu, Dave Touchette. Quantum Log-Approximate-Rank Conjecture is also False. Electronic Colloquium on Computational Complexity (ECCC) TR18-201, 2018.

[6] Debajyoti Bera: Two-sided Quantum Amplitude Amplification and Exact-Error Algorithms. CoRR abs/1605.01828, 2016.

[7] Mark Braverman, Omri Weinstein: A Discrepancy Lower Bound for Information Complexity. Algorithmica 76(3): pp. 846-864, 2016.

[8] H. Buhrman, R. Cleve, J. Watrous, R. de Wolf. Quantum fingerprinting. In Physical Review Letters, 87(16), September 26, 2001.

[9] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In Proceedings of 30th STOC, pp. 63-68, 1998.

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

[11] Arkadev Chattopadhyay, Nikhil S. Mande, Suhail Sherif. The Log-Approximate-Rank Conjecture is False. Electronic Colloquium on Computational Complexity (ECCC) TR18-176, 2018.

[12] J. Forster, M. Krause, S. V. Lokam, R. Mubarakzjanov, N. Schmitt, and H. U. Simon. Relations between communication complexity, linear arrangements, and computational complexity. FSTTCS: Foundations of Software Technology and Theoretical Computer Science, 21, pp. 171 - 182, 2001.

[13] D. Gavinsky and S. Lovett, En route to the log-rank conjecture: New reductions and equivalent formulations, In Proceedings of ICALP 2014, pp. 514–524.

[14] Mika Göös, Toniann Pitassi, Thomas Watson: Deterministic Communication vs. Partition Number. In Proceedings of FOCS 2015, pp. 1077-1088.

[15] Lov K. Grover. A fast quantum mechanical algorithm for database search. In Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, pp. 212–219, 1996.

[16] Peter Høyer and Ronald de Wolf. Improved quantum communication complexity bounds for disjointness and equality. In Proceedings of STACS 2002, pp. 299–310.

[17] B. Kalyanasundaram and G. Schnitger. The Probabilistic Communication Complexity of Set Intersection. SIAM J Discrete Mathematics 5 (4), 1992, pp. 545–557.

[18] Hartmut Klauck: Lower Bounds for Quantum Communication Complexity. SIAM J. Comput. 37(1), pp. 20-46, 2007.

[19] M. Krause. Geometric arguments yield better bounds for threshold circuits and distributed computing. Theoretical Computer Science, 156, pp. 99-117, 1996.

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

[21] Troy Lee, Adi Shraibman: Lower Bounds in Communication Complexity. Foundations and Trends in Theoretical Computer Science 3(4) pp. 263-398 (2009).

[22] Troy Lee, Adi Shraibman: An Approximation Algorithm for Approximation Rank. IEEE Conference on Computational Complexity 2009, pp. 351-357.

[23] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman. Complexity measures of sign matrices. Combinatorica, 27(4) pp. 439-463, 2007.

[24] N. Linial, A. Shraibman: Learning Complexity vs Communication Complexity. Combinatorics, Probability and Computing 18(1-2), pp. 227-245, 2009.

[25] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. Random Structures and Algorithms, 34 pp. 368-394, 2009.

[26] L. Lovász and M. Saks. Mbius functions and communication complexity. In Proceedings of the 29th IEEE Symposium on Foundations of Computer Science, pp. 81-90. IEEE, 1988.

[27] Shachar Lovett. Communication is Bounded by Root of Rank. J. ACM 63(1) pp. 1:1-1:9 (2016)

[28] Shachar Lovett. Personal communication, 2018.

[29] K. Mehlhorn and E. Schmidt. Las Vegas is better than determinism in VLSI and distributed computing. In Proceedings of the 14th ACM Symposium on the Theory of Computing, pp. 330-337, 1982.

[30] M. Mosca G. Brassard, P. Høyer and A. Tapp. Quantum amplitude amplification and estimation. In Quantum Computation and Quantum Information: A Millennium Volume, volume 305 of AMS Contemporary Mathematics Series. American Mathematical Society, 2002.

[31] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[32] I. Newman. Private versus common random bits in communication complexity. Information Processing Letters, 39 pp. 67-71, 1991.

[33] N. Nisan and A. Wigderson. A note on rank vs. communication complexity. Combinatorica, 15(4) pp. 557-566, 1995.

[34] R. Paturi and J. Simon. Probabilistic communication complexity. Journal of Computer and System Sciences, 33(1) pp. 106-123, 1986.

[35] A. Razborov. On the Distributional Complexity of Disjointness. Theoretical ComputerScience 106(2), pp. 385-390, 1992.

[36] A. Razborov. Quantum communication complexity of symmetric predicates. Izvestiya: Mathematics, 67(1) pp. 145–159, 2003.

[37] Thomas Rothvoß. A direct proof for Lovett's bound on the communication complexity of low rank matrices. *CoRR*, abs/1409.6366, 2014.

[38] Makrand Sinha, Ronald de Wolf. Exponential Separation between Quantum Communication and Logarithm of Approximate Rank. Electronic Colloquium on Computational Complexity (ECCC) TR18-204, 2018.

[39] Shengyu Zhang. Efficient quantum protocols for XOR functions. In Proceedings of SODA 2014, pp. 1878-1885.