Lower Bounds in Communication Complexity Based on Factorization Norms

Nati Linial
School of Computer Science and Engineering
Hebrew University
Jerusalem, Israel
nati@cs.huji.ac.il

Adi Shraibman
School of Computer Science and Engineering
Hebrew University
Jerusalem, Israel
adidan@cs.huji.ac.il

ABSTRACT

We introduce a new method to derive lower bounds on randomized and quantum communication complexity. Our method is based on factorization norms, a notion from Banach Space theory. This approach gives us access to several powerful tools from this area such as normed spaces duality and Grothendiek's inequality. This extends the arsenal of methods for deriving lower bounds in communication complexity.

As we show, our method subsumes most of the previously known general approaches to lower bounds on communication complexity. Moreover, we extend all (but one) of these lower bounds to the realm of quantum communication complexity with entanglement.

Our results also shed some light on the question how much communication can be saved by using entanglement. It is known that entanglement can save one of every two qubits, and examples for which this is tight are also known. It follows from our results that this bound on the saving in communication is tight almost always.

Categories and Subject Descriptors: F.2.3 [Analysis of algorithms and problem complexity]: Tradeoffs between Complexity Measures.

General Terms: Theory.

Keywords: Communication complexity, Factorization norms, Discrepancy, Fourier analysis.

1. INTRODUCTION

We study lower bounds for randomized and quantum communication complexity. Our bounds are expressed in terms of *factorization norms*, a concept of great interest in Banach Space Theory which we now introduce. Consider a matrix M as a linear operator between two normed spaces $M : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$. We define its *operator norm* $\|M\|_{\|\cdot\|_X \to \|\cdot\|_Y}$ as the supremum of $\|Mx\|_Y$ over all $x \in X$ with $\|x\|_X = 1$. *Factorization norms*, and in particular the γ_2 norm are defined by considering all possible ways of expressing M as the composition of two linear operators via a given middle normed space. Specifically, the γ_2 norm of an $m \times n$ real matrix B is defined via: [1]

$$\gamma_2(B) = \min_{XY=B} \|X\|_{\ell_2 \to \ell_\infty^m} \|Y\|_{\ell_1^n \to \ell_2}. \quad (1)$$

We introduce here a variation on this definition that plays a key role in our paper. Let A be a sign matrix and let $\alpha \geq 1$

$$\gamma_2^\alpha(A) = \min \gamma_2(B), \quad (2)$$

where the minimum is over all matrices B such that $1 \leq a_{ij} b_{ij} \leq \alpha$ for all i, j. In particular

$$\gamma_2^\infty(A) = \min_{B:\ \forall i,j\ 1 \leq a_{ij} b_{ij}} \gamma_2(B).$$

Let A be a sign matrix and let an error bound $\epsilon > 0$ be given. We consider A's randomized communication complexity and quantum communication complexity with entanglement and denote them by $R_\epsilon(A)$ and $Q_\epsilon^*(A)$ respectively. We are now able to state one of our main theorems:

THEOREM 1. *For every sign matrix A and any $\epsilon > 0$*

$$R_\epsilon(A) \geq 2 \log \gamma_2^{\alpha_\epsilon}(A) - 2 \log \alpha_\epsilon,$$

and

$$Q_\epsilon^*(A) \geq \log \gamma_2^{\alpha_\epsilon}(A) - \log \alpha_\epsilon - 2,$$

where $\alpha_\epsilon = \frac{1}{1-2\epsilon}$. Both bounds are tight up to the additive term.

These bounds are proved in Sections 3.1 and 3.2. Although the two proofs are rather different, they both rely on the key observation that γ_2 and its variants are complexity measures of matrices. It is this basic idea and its broad applicability that we consider as the key contributions of our work.

The usefulness of the lower bounds in Theorem 1 is further elaborated in Section 4. There we prove that these bounds extend and improve previously known general bounds on randomized and quantum communication complexity. It is shown that our bounds extend the discrepancy method initiated in [17, 1]. It also extends a general bound in terms of the trace norm from [16], and bounds using the Fourier Transform of boolean functions studied in [15, 6]. (Some of the basic features of these methods are explained in Section 4). We are also able to generalize other bounds, in terms

[1] In order to develop some intuition for this definition, it is useful to observe that $\|Y\|_{\ell_1^n \to \ell_2}$ is the largest ℓ_2 norm of a column of Y, and $\|X\|_{\ell_2 \to \ell_\infty^m}$ is the largest ℓ_2 norm of a row of X.

STOC'07, June 11–13, 2007, San Diego, California, USA.

of singular values, and entropy, proved in [6]. Thus, our work immediately yields simpler and more transparent proofs of previously known bounds. It also implies that bounds based on discrepancy arguments and on Fourier analysis apply to quantum communication complexity with entanglement, thus answering a well-known open question in that area.

In Section 3.5 we prove an *upper bound* on communication complexity in terms of factorization norms.

CLAIM 2. *The one round probabilistic communication complexity with public random bits of a matrix A is at most $O((\gamma_2^\infty(A))^2)$. The bound is tight.*

We raise the possibility that a better bound may hold in which $\gamma_2^\infty(A)$ is replaced by γ_2^α for some small α.

Another intriguing open question is whether $R_\epsilon(A) \geq \Omega(\log \gamma_2)$ for every sign matrix A. We are able to show that if $\gamma_2(A) \geq \Omega(\sqrt{n})$ (a condition satisfied by almost all $n \times n$ sign matrices), then indeed $R_\epsilon(A), Q_\epsilon^*(A) \geq \Omega(\log n)$.

In Section 5 we consider two specific families of functions. We estimate the value of different complexity measures considered in this paper for these families.

2. BACKGROUND AND NOTATIONS

Factorization norms.

We have already introduced the definition of the factorization norm γ_2 and its variations γ_2^α. We next collect several basic properties of these parameters

PROPOSITION 3. *For every $m \times n$ sign matrix A and every $\alpha \geq 1$,*

1. $\gamma_2^\infty \leq \gamma_2^\alpha(A) \leq \gamma_2(A) \leq \sqrt{rank(A)}$.
2. *$\gamma_2^\alpha(A)$ is a decreasing, convex function of α.*
3. *It is possible to express $\gamma_2^\alpha(A)$ as the optimum of a semidefinite program of size $O(mn)$.*

The first statement is proved in [10], the second is not hard, and the third is proved in Section 3.3.

Fourier analysis - some basics.

Identify $\{0,1\}^n$ with $\mathbb{Z}_2^n$. For functions $f, g : \{0,1\}^n \to \mathbb{R}$, define

$$\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \mathbb{Z}_2^n} f(x) \cdot g(x),$$

and $\|f\|_2 = \sqrt{\langle f, f \rangle}$. Corresponding to every $z \in \mathbb{Z}_2^n$, is a character of $\mathbb{Z}_2^n$ denoted χ_z

$$\chi_z(x) = (-1)^{\langle z, x \rangle}.$$

The Fourier coefficients of f are $\hat{f}_z = \langle f, \chi_z \rangle$ for all $z \in \mathbb{Z}_2^n$. For $M = 2^m$ and $N = 2^n$, we occasionally consider a real $M \times N$ matrix B as a function from $\mathbb{Z}_2^m \times \mathbb{Z}_2^n$ to $\mathbb{R}$. Thus the (i,j)-entry of B, B_{ij}, is also denoted $B_{z,z'}$, where z and z' are the binary representations of i and j respectively. For B as above and $(z, z') \in \mathbb{Z}_2^m \times \mathbb{Z}_2^n$ we denote the corresponding Fourier coefficient of B (thought of as a function) by $\hat{B}_{z,z'}$.

We use the following observation in our proofs:

OBSERVATION 4. *Let $B = xy^t$ be a $2^m \times 2^n$ sign matrix of rank 1. Then $\hat{B}_{z,z'} = \hat{x}_z \cdot \hat{y}_{z'}$ for all $z \in \mathbb{Z}_2^m$ and $z' \in \mathbb{Z}_2^n$. Here x and y are viewed as real functions on $\mathbb{Z}_2^m$ resp. $\mathbb{Z}_2^n$.*

Additional notations.

Let A and B be two real matrices. We use the following notations:

- $s_1(B) \geq s_2(B) \geq \ldots \geq 0$ are the singular values of B.
- $\|B\|_1 = \sum |b_{ij}|$ is its ℓ_1 norm, $\|B\|_2 = \sqrt{\sum b_{ij}^2}$ is its ℓ_2 (Frobenius) norm, and $\|B\|_\infty = \max_{ij} |b_{ij}|$ is its ℓ_∞ norm.
- The inner product of A and B is denoted $\langle A, B \rangle = \sum_{ij} a_{ij} b_{ij}$.

We should note a difference between our corresponding definitions for matrices and for boolean functions. In the latter case, the inner product $\langle \cdot, \cdot \rangle$, and the ℓ_2 norm $\| \cdot \|_2$, are normalized.

2.1 Background on Grothendieck's Inequality

We review Grothendieck's inequality which is an important tool in our proofs, see e.g. [14, pg. 64].

THEOREM 5 (GROTHENDIECK'S INEQUALITY). *There is a universal constant $1.5 \leq K_G \leq 1.8$ such that for every real matrix B and every $k \geq 1$*

$$\max \sum b_{ij} \langle u_i, v_j \rangle \leq K_G \max \sum b_{ij} \epsilon_i \delta_j. \tag{3}$$

where the max are over the choice of $u_1, \ldots, u_m, v_1, \ldots, v_n$ as unit vectors in $\mathbb{R}^k$ and $\epsilon_1, \ldots, \epsilon_m, \delta_1, \ldots, \delta_n \in \{\pm 1\}$.

We denote by γ_2^* the dual norm of γ_2, i.e. for every real matrix B

$$\gamma_2^*(B) = \max_{C : \gamma_2(C) \leq 1} \langle B, C \rangle.$$

We note that for any real matrix γ_2^* and $\| \cdot \|_{\infty \to 1}$ are equivalent up to a small multiplicative factor, viz.

$$\|B\|_{\infty \to 1} \leq \gamma_2^*(B) \leq K_G \|B\|_{\infty \to 1}. \tag{4}$$

The left inequality is easy, and the right inequality is a reformulation of Grothendieck's inequality. Both use the observation that the left hand side of (3) equals $\gamma_2^*(B)$, and the max term on the right hand side is $\|B\|_{\infty \to 1}$.
Additional useful corollaries of Grothendieck's inequality are collected below.

LEMMA 6. *Every real matrix B can be expressed as $B = \sum_i w_i x_i y_i^t$, where $w_1, \ldots, w_s$ are positive reals, and $x_1, \ldots, x_s$, $y_1, \ldots, y_s$ are sign vectors such that*

$$\gamma_2(B) \leq \sum_i w_i \leq K_G \cdot \gamma_2(B). \tag{5}$$

PROOF. We recall ν, the *nuclear norm* from l_1 to l_∞ of a real matrix B, that is defined as follows. $\nu(B) = \min \sum |w_i|$ such that B can be expressed as $\sum w_i x_i y_i^t = B$ for some choice of sign vectors $x_1, x_2, \ldots, y_1, y_2 \ldots$. It is known that ν is the norm dual to $\| \cdot \|_{\infty \to 1}$. See [5] for more details.

It is a simple consequence of the definition of duality and (4) that for every real matrix B

$$\gamma_2(B) \leq \nu(B) \leq K_G \cdot \gamma_2(B). \tag{6}$$

The claim follows now if we note that in the definition of $\nu(B)$ the w_i can be made positive, by replacing the appropriate x_i by $-x_i$. □

The following corollary is a simple consequence of Lemma 6.

COROLLARY 7. *Let B be a real matrix satisfying $\gamma_2(B) \leq 1$. Then for every $\delta > 0$ there are sign vectors $\phi_1, \phi_2 \ldots, \psi_1, \psi_2 \ldots \in \{\pm 1\}^k$ for some integer k such that*

$$\frac{b_{ij}}{K_G} - \delta \leq \frac{1}{k}\langle \phi_i, \psi_j \rangle \leq b_{ij} + \delta, \tag{7}$$

for all i, j.

PROOF. Let $M = \frac{1}{K_G}B$. By Inequality (6), $\nu(M) \leq 1$. Consider the expansion $M = \sum w_i x_i y_i^t$ with $w_i > 0$ for which $\nu(M) = \sum w_i$. If the w_i happen to be rational, say $w_i = \frac{u_i}{k}$ (k is the common denominator), then we can satisfy the claim with $\delta = 0$. Construct sign matrices P, Q that have u_i columns (rows) equal to x_i (resp. y_i) in this order. Then $\frac{B}{K_G} = M = \frac{1}{k}PQ$. The claim follows with ϕ_i, ψ_j being the rows (columns) of P and Q respectively. The general case follows by approximating the w_i's by rationals. □

3. A NEW LOWER BOUND TECHNIQUE IN COMMUNICATION COMPLEXITY

Let us recall some terminology:

- The deterministic communication complexity of a sign matrix A is denoted by $CC(A)$.
- Its quantum communication complexity is $Q_\epsilon(A)$. When prior entanglement is allowed we denote it by $Q^*_\epsilon(A)$.
- The randomized communication complexity is $R_\epsilon(A)$.

In the latter two definitions ϵ is the error bound. Since the value of ϵ is usually immaterial, we simply omit it whenever this causes no confusion. That the value of ϵ is inconsequential follows from a simple amplification-by-repetition argument (e.g. [9]). For illustration, this argument yields e.g., $Q^*_\epsilon(A) \leq O(Q^*_{1/3}(A) \cdot \log \frac{1}{\epsilon})$ for every sign matrix A and any $\epsilon > 0$. When there is no mention of ϵ it is assumed to be 1/3.

In this section we review some of the basic ideas in the field and prove our results. In Section 4 we compare our bounds with previously known bounds.

We should note first, that a basic observation underlying our new bounds is that γ_2 is a complexity measure for matrices, in the same way that the *rank* has long been used (explicitly or implicitly) as a measure of complexity for matrices. For a more elaborate discussion on this subject, see [10].

3.1 Randomized communication complexity

In order to find lower bounds on randomized communication complexity, one uses the following observation

OBSERVATION 8. *A sign matrix A satisfies $R_\epsilon(A) \leq c$ if and only if there are sign matrices D_i, $i = 1, \ldots, m$, satisfying $CC(D_i) \leq c$ and a probability distribution $(p_1, \ldots, p_m)$ such that*

$$\|A - \sum_{i=1}^{m} p_i D_i\|_\infty \leq 2\epsilon. \tag{8}$$

Condition (8) can be combined with the fact that each of the matrices D_i can be partitioned into at most 2^c monochromatic rectangles. These two facts are used by the discrepancy method to derive a lower bound on $R_\epsilon(A)$.

There is an alternative route (see [15]) that proceeds from here using Fourier analysis.

As we observe next, $\gamma_2^\alpha(A)$ fits very well into this general frame.

THEOREM 9. *For every sign matrix A and any $\epsilon > 0$*

$$R_\epsilon(A) \geq 2 \log \gamma_2^{\alpha_\epsilon}(A) - 2 \log \alpha_\epsilon,$$

where $\alpha_\epsilon = \frac{1}{1-2\epsilon}$.

PROOF. Let D_i, $i = 1, \ldots, m$, and p be as above, and denote $B = \frac{1}{1-2\epsilon}\sum_{i=1}^m p_i D_i$. Recall that $\log(rank(A)) \leq CC(A)$ for every sign matrix A. Thus, for every $i = 1, \ldots, m$

$$\gamma_2(D_i) \leq (rank(D_i))^{1/2} \leq 2^{CC(D_i)/2} \leq 2^{R_\epsilon(A)/2}.$$

The first inequality is from Proposition 3. Since γ_2 is a norm

$$\begin{aligned} \gamma_2(B) &= \frac{1}{1-2\epsilon}\gamma_2(\sum_{i=1}^m p_i D_i) \\ &\leq \frac{1}{1-2\epsilon}\sum_{i=1}^m p_i \gamma_2(D_i) \leq \frac{1}{1-2\epsilon}2^{R_\epsilon(A)/2}. \end{aligned}$$

On the other hand it follows from Equation (8) that $1 \leq a_{ij}b_{ij} \leq \frac{1}{1-2\epsilon}$. Hence, by the definition of γ_2^α (Equation (2)), for $\alpha = \frac{1}{1-2\epsilon}$

$$\gamma_2^\alpha(A) \leq \gamma_2(B) \leq \frac{1}{1-2\epsilon}2^{R_\epsilon(A)/2}.$$

□

3.2 Quantum communication complexity

A possible first step in search of lower bounds in quantum communication complexity is the following fact, variants of which were observed by several authors [16, 18, 3, 7].

LEMMA 10. *Given a sign matrix A, let $P = (p_{ij})$ be the acceptance probabilities of a quantum protocol for A with complexity C. Then there are matrices X, Y such that $P = XY$ and*

$$\|X\|_{2\to\infty}, \|Y\|_{1\to 2} \leq 2^{C/2}. \tag{9}$$

If prior entanglement is not used, then the matrices X and Y in Condition (9) can be chosen to have rank at most 2^{2C}.

As mentioned, there are several similar statements in the literature, but we could not find a reference for this precise statement. We defer a proof of Lemma 10 to a full version of this paper. When there is no prior entanglement, lemma 10 yields a condition analogous to observation 8 and then bounds via discrepancy and Fourier analysis can be likewise derived. However, this was not known for the model of quantum communication complexity with entanglement. Our method provides a coherent way to extend previously known bounds (based on the discrepancy and Fourier transform methods) for the model allowing entanglement. The next theorem uses Lemma 10 to give a bound on quantum communication complexity in terms of γ_2^α.

THEOREM 11. *For every sign matrix A and any $\epsilon > 0$*

$$Q^*_\epsilon(A) \geq \log \gamma_2^{\alpha_\epsilon}(A) - \log \alpha_\epsilon - 2,$$

where $\alpha_\epsilon = \frac{1}{1-2\epsilon}$.

PROOF. Let $P = (p_{ij})$ be the acceptance probabilities of an optimal quantum protocol for A. By Lemma 10, $\gamma_2(P) \leq 2^{Q^*_\epsilon(A)}$.

On the other hand, by definition, $p_{ij} \leq \epsilon$ when $a_{ij} = -1$ and $p_{ij} \geq 1-\epsilon$ when $a_{ij} = 1$. Thus, if we let $B = \frac{1}{1-2\epsilon}(2P - J)$, we get that $b_{ij}a_{ij} \geq 1$ for all i, j and

$$\begin{aligned}\gamma_2(B) &= \gamma_2(\frac{1}{1-2\epsilon}(2P-J)) \leq \frac{1}{1-2\epsilon}(2\gamma_2(P)+1)\\ &\leq \frac{1}{1-2\epsilon}\left(2^{Q^*_\epsilon(A)+2}\right).\end{aligned}$$

We conclude that

$$\gamma_2^{\alpha_\epsilon}(A) \leq \gamma_2(B) \leq \frac{1}{1-2\epsilon}2^{Q^*_\epsilon(A)+2},$$

and hence

$$Q^*_\epsilon(A) \geq \log \gamma_2^{\alpha_\epsilon}(A) - \log \alpha_\epsilon - 2,$$

for $\alpha_\epsilon = \frac{1}{1-2\epsilon}$. □

3.3 Employing duality

One interesting aspect of our main result is that it improves several previously known bounds. This point is elaborated on in Section 4. Another noteworthy point is that our bounds are expressed in terms of $\gamma_2^\alpha(\cdot)$, a quantity that can be efficiently computed using SDP. A particularly useful consequence of this observation is that *SDP duality* makes it often possible to derive good (sometimes even optimal) lower bounds on communication complexity.

The following theorem gives an expression for γ_2^α that is derived using SDP duality.

THEOREM 12. *For every sign matrix A and $\alpha \geq 1$*

$$\begin{aligned}\gamma_2^\alpha(A)^{-1} = \quad &\min \quad \gamma_2^*((P-Q)\circ A)\\ &s.t. \quad P, Q \geq 0\\ &\quad \sum p_{ij} - \alpha q_{ij} = 1,\end{aligned}$$

and also

$$\begin{aligned}\gamma_2^\alpha(A) = \max \quad &\langle A, B\rangle - (\alpha-1)\textstyle\sum_{ij:a_{ij}\neq sign(b_{ij})}|b_{ij}|\\ s.t. \quad &\gamma_2^*(B) = 1\end{aligned}.$$

In particular, for $\alpha = \infty$

$$\begin{aligned}\gamma_2^\infty(A)^{-1} = \quad &\min \quad \gamma_2^*(P\circ A)\\ &s.t. \quad P \geq 0\\ &\quad \sum p_{ij} = 1,\end{aligned}$$

and also

$$\begin{aligned}\gamma_2^\infty(A) = \quad &\max \quad \langle A, B\rangle\\ &s.t. \quad sign(B) = A \text{ and } \gamma_2^*(B) = 1.\end{aligned}$$

PROOF. We start by showing that for every sign matrix A and $\alpha > 1$

$$\begin{aligned}\gamma_2^\alpha(A)^{-1} = \quad &\max \mu\\ \text{s.t. for all } i,j \quad &\mu \leq a_{ij}b_{ij} \leq \alpha\mu \qquad (10)\\ &\gamma_2(B) \leq 1.\end{aligned}$$

Denote by $\mu(A)$ the maximum on the right hand side above. Let C be a matrix such that $\gamma_2(C) = \gamma_2^\alpha(A)$ and $1 \leq a_{ij}c_{ij} \leq \alpha$, and take $B = \gamma_2^\alpha(A)^{-1}C$. Then, $\gamma_2(B) \leq 1$ and $\gamma_2^\alpha(A)^{-1} \leq a_{ij}b_{ij} \leq \alpha\gamma_2^\alpha(A)^{-1}$, implying that $\mu(A) \geq \gamma_2^\alpha(A)^{-1}$. To prove the inverse inequality, let B be a matrix such that $\gamma_2(B) \leq 1$ and $\mu(A) \leq a_{ij}b_{ij} \leq \alpha\mu(A)$, and take $C = \mu(A)^{-1}B$. Then $1 \leq a_{ij}c_{ij} \leq \alpha$ and $\gamma_2(C) \leq \mu(A)^{-1}$, implying that $\gamma_2^\alpha \leq \mu(A)^{-1}$ or equivalently $\mu(A) \leq \gamma_2^\alpha(A)^{-1}$.

Note that (10) is a semidefinite program, since the condition $\gamma_2(B) \leq 1$ is expressible as an SDP. By SDP duality

$$\begin{aligned}\gamma_2^\alpha(A)^{-1} = \quad &\min \quad \gamma_2^*((P-Q)\circ A)\\ &\text{s.t.} \quad P, Q \geq 0 \qquad (11)\\ &\quad \sum p_{ij} - \alpha q_{ij} = 1,\end{aligned}$$

proving the first identity. We use this to prove the second identity, i.e. that

$$\begin{aligned}\gamma_2^\alpha(A) = \max \quad &\langle A, B\rangle - (\alpha-1)\textstyle\sum_{ij:a_{ij}\neq sign(b_{ij})}|b_{ij}|\\ \text{s.t.} \quad &\gamma_2^*(B) = 1\end{aligned}.$$

To see that the optimum of the above SDP is indeed equal to $\gamma_2^\alpha(A)$, note that by choosing B such that $P - Q = B \circ A$, the SDP in (11) is equivalent to

$$\begin{aligned}\min \quad &\gamma_2^*(B)\\ \text{s.t.} \quad &\textstyle\sum_{ij:a_{ij}=sign(b_{ij})}|b_{ij}| - \alpha\sum_{ij:a_{ij}\neq sign(b_{ij})}|b_{ij}| = 1\end{aligned}.$$

Since both $\gamma_2^*(B)$ and the constraints above are homogeneous in B, the optimum of this SDP is the inverse of

$$\begin{aligned}\max \quad &\langle A, B\rangle - (\alpha-1)\textstyle\sum_{ij:a_{ij}\neq sign(b_{ij})}|b_{ij}|\\ \text{s.t.} \quad &\gamma_2^*(B) = 1\end{aligned},$$

as required.

The statements regarding γ_2^∞ follow by considering the corresponding expressions for γ_2^α and taking α to infinity. □

As usual, the advantage of this result is that any feasible solution to the SDPs in Theorem 12 yields a lower bound for $\gamma_2^\alpha(A)$ or $\gamma_2^\infty(A)$. What is left is to find good feasible solutions.

Also note that by Grothendieck's inequality (Theorem 5, and Inequality (4)), we can replace γ_2^* with $\|\cdot\|_{\infty\to 1}$ in Theorem 12, without changing the value of the SDPs by more than a factor of K_G.

3.4 How does $\log\gamma_2$ fit in?

As we just saw, randomized and quantum communication complexity are bounded below by $\log\gamma_2^\alpha$. It is an interesting open question how these two parameters compare with $\log\gamma_2$. For most $m \times n$ sign matrices A with $m \geq n$, it holds that

1. $\gamma_2(A) = \Theta(\sqrt{n})$,
2. $R_\epsilon(A) = \log n - O_\epsilon(1)$,
3. $Q_\epsilon(A) = \frac{1}{2}\log n - O_\epsilon(1)$.

The first item was shown in [10], alongside the fact that $\gamma_2^\infty(A) = \Theta(\sqrt{n})$ for random matrices. The other two items follow therefore, from Theorems 9 and 11. As shown by the next claim, the first condition implies the other two.

CLAIM 13. *Let A be an $m \times n$ sign matrix with $m \geq n$. If $\gamma_2(A) \geq \Omega(\sqrt{n})$, then $R(A) \geq \log n - O(1)$, and $Q^*(A) \geq \frac{1}{2}\log n - O(1)$.*

This claim is an easy consequence of the following lemma

LEMMA 14. *Let A be an $m \times n$ sign matrix with $m \geq n$. Then for every $\delta > 0$,*

$$\gamma_2(A) \leq \gamma_2^{1+\delta}(A) + \frac{\delta}{2}(\sqrt{n}+1). \quad (12)$$

PROOF. Let B be a matrix with $1 \leq a_{ij}b_{ij} \leq 1+\delta$ and $\gamma_2(B) = \gamma_2^{1+\delta}(A)$. Since γ_2 is a norm, we may write

$$\gamma_2(A) \leq \gamma_2(B - \frac{\delta}{2}J) + \gamma_2(B - \frac{\delta}{2}J - A).$$

Since all elements of the matrix $B - \frac{\delta}{2}J - A$ have absolute value $\leq \frac{\delta}{2}$, the claim follows using linearity of the norm, the fact that $\gamma_2 \leq \min\{\sqrt{m}, \sqrt{n}\}$ for every $m \times n$ sign matrix (which follows from the trivial factorizations $A \cdot I = A$ resp. $I \cdot A = A$), and that $\gamma_2(J) = 1$. □

It is now a simple matter to prove Claim 13. If $\gamma_2(A) \geq c\sqrt{n}$, then $\gamma_2^{1+c}(A) > \frac{c}{2}(\sqrt{n}-1)$ from which the Claim follows, by Theorem 1.

We cannot rule out the intriguing possibility that Claim 13 is a tip of something bigger and that R_ϵ as well as Q_ϵ^* are in fact polynomially equivalent to $\log \gamma_2$. This point is discussed further in Section 6.

3.5 An upper bound in terms of γ_2^∞

We have established so far lower bounds on communication complexity in terms of γ_2^α. Here we show an *upper bound* that is "only" exponentially larger than these lower bounds, in terms of γ_2^∞. We also observe that this bound is essentially tight, if we insist on using γ_2^∞. It is not impossible that better bounds exist which are expressed in terms of γ_2^α with finite α. The idea behind Claim 15 is not new, e.g. [8], and is included for completeness sake.

CLAIM 15. *The one round probabilistic communication complexity (with public random bits) of a matrix A is at most $O((\gamma_2^\infty(A))^2)$.*

This bound is tight up to the (second) power of $\gamma_2^\infty(A)$. This is illustrated by the matrix D_k that corresponds to the disjointness function on k bits, as seen in Section 5.1.

PROOF OF CLAIM 15. Let x be a vector of length k and let T be a multiset with elements in $[k]$. We denote by $x|_T$ the restriction of x to the coordinates indexed by the elements of T. For example if $x = (10, 1, 17, 42, 8)$ and $T = (1, 2, 2, 5)$, then $x|_T = (10, 1, 1, 8)$. The communication protocol we consider is as follows: Let B be a real matrix satisfying $\gamma_2(B) = \gamma_2^\infty(A)$ and $1 \leq b_{ij}a_{ij}$ for all i, j. By Corollary 7 there are sign vectors $x_1, \ldots, x_m, y_1, \ldots, y_n \in \{\pm 1\}^k$ for some $k \geq 1$ such that

$$\frac{b_{ij}}{K_G\gamma_2(B)} \leq \frac{1}{k}\langle x_i, y_j\rangle \leq \frac{b_{ij}}{\gamma_2(B)}. \quad (13)$$

for all i, j.

Given indices i and j, the row player uses the publicly available random bits to select at random a multiset T with elements from $[k]$. He sends $x_i|_T$ to the column player who then computes $\langle x_i|_T, y_j|_T\rangle$ and outputs the sign of the result. Next we analyze the complexity and the error probability of this protocol.

Let $\mu > 0$ and consider two sign vectors x and y of length k, such that $|\langle x, y\rangle| \geq \mu k$. We wish to bound the probability that for a random multiset T of size K with elements from $[k]$, $sign(\langle x, y\rangle) \neq sign(\langle x|_T, y|_T\rangle)$. Assume w.l.o.g. that $x = (1, 1, \ldots, 1)$ and that $\langle x, y\rangle > 0$. Denote the number of -1s in y by Qk, where by our assumptions $Q \leq \frac{1-\mu}{2}$. We should bound the probability that y_T contains at least $K/2$ -1's for a random multiset T of size K. This is exactly the probability of picking more -1's than 1's when we sample independently K random bits each of which is -1 (resp. 1) with probability Q (resp. $1-Q$.) By Chernoff bound the probability of this event is at most:

$$e^{-2(1/2-Q)^2K} \leq e^{-K\mu^2/2}.$$

Thus, to achieve a constant probability of error it is enough to take $K = O(\mu^{-2})$. By Equation (13), $|\langle x_i, y_j\rangle| \geq \frac{k}{K_G\gamma_2(B)}$, thus the complexity of our protocol (with constant probability of error) is at most $O((\gamma_2(B))^2) = O((\gamma_2^\infty(A))^2)$. □

4. RELATIONS WITH OTHER BOUNDS

We prove next that the bounds in Theorems 9 and 11 nicely generalize some of the previously known bounds for communication complexity. In Section 4.1 we consider the discrepancy method and in Section 4.2 bounds involving singular values (Ky Fan norms and in particular the trace norm, are discussed). In Sections 4.3 and 4.3.2 lower bounds that are based on Fourier analysis of boolean functions are examined, and in Section 4.4, bounds in terms of entropy.

4.1 The discrepancy method

Let A be a sign matrix, and let P be a probability measure on the entries of A. The P-discrepancy of A, denoted $disc_P(A)$, is defined as the maximum over all combinatorial rectangles R in A of $|P^+(R) - P^-(R)|$, where P^+ [P^-] is the P-measure of the positive entries [negative entries]. The discrepancy of a sign matrix A, denoted $disc(A)$, is the minimum of $disc_P(A)$ over all probability measures P on the entries of A.

The discrepancy method, introduced in [17, 1], was the first general method for deriving lower bounds for randomized communication complexity. It is based on the following fact: for every sign matrix A

$$Q_\epsilon(A), R_\epsilon(A) \geq \Omega\left(\log\left(\frac{1-2\epsilon}{disc(A)}\right)\right).$$

See [9] for a more elaborate discussion on this bound for randomized communication complexity, and [7] for the first proof extending this bound to the realm of quantum communication complexity.

The following theorem was proved in [11] [2]

THEOREM 16. *For every sign matrix A*

$$\frac{1}{8}\gamma_2^\infty(A) \leq disc(A)^{-1} \leq 8\gamma_2^\infty(A).$$

An immediate corollary of Theorem 16 and Theorems 9 and 11 is the following.

THEOREM 17. *For every sign matrix A and any $\epsilon > 0$*

$$R_\epsilon(A) \geq 2\log\left(\frac{1-2\epsilon}{disc(A)}\right) - O(1),$$

[2] As observed in [11], γ_2^∞ is the same as *margin complexity*, a parameter of interest in the field of machine learning.

and

$$Q^*_\epsilon(A) \geq \log\left(\frac{1-2\epsilon}{disc(A)}\right) - O(1).$$

Both bounds are tight up to the additive term.

This settles the widely known open question whether the discrepancy bound holds for quantum communication complexity with entanglement.

Our bounds are it terms of γ_2^α, and as mentioned above, γ_2^∞ (which is smaller than γ_2^α) is equal up to a multiplicative constant to the inverse of discrepancy. In Section 5.1 we show an example where γ_2^∞ is significantly smaller than γ_2^α for small α. The behavior of γ_2^α as a function of α is an interesting subject for research, as further discussed in Section 5.1 and Section 6.

4.2 Bounds involving singular values

4.2.1 The trace norm

We recall that the *trace norm* $\|A\|_{tr}$ of a real matrix A is the sum of its singular values. We introduce the following concept (from [16]), analogous to γ_2^α:

$$\|A\|_{tr}^\alpha = \min\left\{\|B\|_{tr} : 1 \leq a_{ij}b_{ij} \leq \alpha\right\}.$$

The following bound on Q^*_ϵ was proved in [16].

THEOREM 18. *For every $n \times n$ sign matrix A and any $\epsilon > 0$, let $\alpha_\epsilon = \frac{1}{1-2\epsilon}$, then*

$$Q^*_\epsilon(A) \geq \Omega(\log(\|A\|_{tr}^{\alpha_\epsilon}/n)).$$

The trace norm and γ_2 are related by the following inequality. For every real $m \times n$ matrix A

$$\|A\|_{tr} \leq \sqrt{mn} \cdot \gamma_2(A). \tag{14}$$

See e.g. [10, Sec. 3] for a proof. It should be clear then, that $\|A\|_{tr}^\alpha \leq \sqrt{mn} \cdot \gamma_2^\alpha(A)$ for every $m \times n$ sign matrix A and every $\alpha \geq 1$.

Therefore, Theorem 18 is a consequence of Theorem 11. Moreover, as shown in Section 5.2, the bound in Theorem 11 can be significantly better than what Theorem 18 yields.

While the bounds in terms of factorization norms are better than those derived from discrepancy and from trace norm, the latter two methods are incomparable. Examples in Section 5.1 and 5.2 demonstrate that the inverse of discrepancy can be much larger than $\|\cdot\|_{tr}^{\alpha_\epsilon}$ and vice versa.

4.2.2 Ky Fan norms

The Ky Fan k-norm of a matrix A which we denote by $\|\cdot\|_\mathcal{K}$ is defined as $\sum_{i=1}^k s_i(A)$, the sum of the k largest singular values of A. Two interesting instances are the Ky Fan n-norm which is the trace norm and the Ky Fan 1-norm - the operator norm from ℓ_2 to ℓ_2.
The following theorem was proved in [6]

THEOREM 19. *[6, th. 6.10] For every $n \times n$ sign matrix A:*
If $\|A\|_\mathcal{K} \geq n\sqrt{k}$, then $Q(f) \geq \Omega(\log(\frac{\|A\|_\mathcal{K}}{n}))$.
If $\|A\|_\mathcal{K} \leq n\sqrt{k}$, then

$$Q(f) \geq \Omega(\log(\frac{\|A\|_\mathcal{K}}{n}))/(\log\sqrt{k} - \log(\frac{\|A\|_\mathcal{K}}{n}) + 1)).$$

We prove

THEOREM 20. *For every $n \times n$ sign matrix A and for every $\delta > 0$*

$$\gamma_2^{1+\delta}(A) \geq \frac{1}{n}\|A\|_\mathcal{K} - \delta \cdot \sqrt{k}$$

PROOF. Let B be a matrix such that $\gamma_2(B) = \gamma_2^{1+\delta}(A)$ and $1 \leq a_{ij}b_{ij} \leq 1+\delta$. By the triangle inequality

$$\|B\|_\mathcal{K} \geq \|A\|_\mathcal{K} - \|A-B\|_\mathcal{K} \geq \|A\|_\mathcal{K} - \delta\sqrt{k}n.$$

To prove the latter inequality, let $M = A - B$ and note that

$$\begin{aligned}\|M\|_\mathcal{K} &= \sum_1^k s_i(M) \leq \sqrt{k}\sqrt{\sum_1^k s_i(M)^2} \\ &\leq \sqrt{k}\sqrt{\sum_1^n s_i(M)^2} = \sqrt{k}\|M\|_2.\end{aligned}$$

The first inequality is Cauchy-Schwartz and the last identity can be found e.g., in [2, p. 7]. It is left to observe that by (14)

$$\|B\|_\mathcal{K} \leq \|B\|_{tr} \leq \gamma_2(B) \cdot n = \gamma_2^{1+\delta}(A) \cdot n.$$

□

Theorems 11 and 20 imply that Klauck's bound holds as well for quantum communication complexity with entanglement

THEOREM 21. *For every $n \times n$ sign matrix A:*
If $\|A\|_\mathcal{K} \geq n\sqrt{k}$, then $Q^*(f) \geq \Omega(\log(\frac{\|A\|_\mathcal{K}}{n}))$.
If $\|A\|_\mathcal{K} \leq n\sqrt{k}$, then $Q^*(f) \geq \Omega(\log(\frac{\|A\|_\mathcal{K}}{n}))/(\log\sqrt{k} - \log(\frac{\|A\|_\mathcal{K}}{n}) + 1))$.

PROOF. If $\|A\|_\mathcal{K} \geq n\sqrt{k}$ then

$$Q^*_{1/6}(A) \geq \log\gamma_2^{3/2}(A) - O(1) \geq \log(\frac{\|A\|_\mathcal{K}}{n}) - O(1).$$

The first inequality is by theorem 11 and the second follows from Theorem 20. Consequently, $Q^*(A) \geq \Omega(Q^*_{1/3}(A)) \geq \Omega(\log(\frac{\|A\|_\mathcal{K}}{n}))$.

If $\|A\|_\mathcal{K} \leq n\sqrt{k}$ take $\epsilon = \frac{\frac{\|A\|_\mathcal{K}}{n\sqrt{k}}}{4+2\frac{\|A\|_\mathcal{K}}{n\sqrt{k}}}$, so that $\alpha_\epsilon = 1 + \frac{\|A\|_\mathcal{K}}{2n\sqrt{k}}$. We have

$$Q^*_\epsilon(A) \geq \log\gamma_2^{\alpha_\epsilon}(A) - O(1) \geq \log(\frac{\|A\|_\mathcal{K}}{n}) - O(1),$$

By amplification of error

$$Q^*(A) \geq \Omega\left(\frac{Q^*_\epsilon(A)}{\log\epsilon^{-1}}\right) \geq \Omega\left(\frac{\log(\frac{\|A\|_\mathcal{K}}{n})}{\log\sqrt{k} - \log(\frac{\|A\|_\mathcal{K}}{n}) + 1}\right).$$

□

4.3 Fourier analysis

We prove here that the bounds on communication complexity in Theorems 9 and 11 subsume previous bounds using Fourier analysis [15, 6] which we review next.

4.3.1 Using the diagonal Fourier coefficients

Any deterministic communication protocol for a sign matrix A naturally partitions it into monochromatic combinatorial rectangles. By Observation 8, if A has randomized

communication complexity at most c then there are rectangles R_i and weights $w_i \in [0,1]$ such that

$$\|A - \sum_i w_i R_i\|_\infty \leq \epsilon,$$

and $\sum_i w_i \leq 2^c$. Raz [15] used this observation and properties of the Fourier transform to derive lower bounds on randomized communication complexity. These ideas were extended by Klauck [6] to quantum communication complexity:

THEOREM 22. *[6, th. 4.1] Let A be a $2^n \times 2^n$ sign matrix. Let E be a set of σ_0 diagonal elements in A and denote $\sigma_1 = \sum_{(z,z)\in E} |\hat{A}_{z,z}|$.*
If $\sigma_1 \geq \sqrt{\sigma_0}$, then $Q(f) \geq \Omega(\log(\sigma_1))$.
If $\sigma_1 \leq \sqrt{\sigma_0}$, then $Q(f) \geq \Omega(\log(\sigma_1)/(\log\sqrt{\sigma_0} - \log\sigma_1 + 1))$.

These bounds can be useful in the study of certain specific matrices. In general, e.g. for random matrices they are rather weak.

Ideas from Raz and Klauck's proofs lead to the following theorem and the conclusion that Theorem 11 yields bounds at least as good as those achieved by Fourier analysis. What is more, this proof technique works as well for quantum communication complexity with prior entanglement.

THEOREM 23. *Let A be a $2^n \times 2^n$ sign matrix, and E be a set of σ_0 diagonal elements with $\sigma_1 = \sum_{(z,z)\in E} |\hat{A}_{z,z}|$. Then $\gamma_2^{1+\delta}(A) \geq \Omega(\sigma_1 - \delta \cdot \sqrt{\sigma_0})$ for every $\delta > 0$.*

A corollary of Theorem 23 and Theorem 11 is

THEOREM 24. *Let A be a $2^n \times 2^n$ sign matrix. Let E be a set of σ_0 diagonal elements in A and denote $\sigma_1 = \sum_{(z,z)\in E} |\hat{A}_{z,z}|$.*
If $\sigma_1 \geq \sqrt{\sigma_0}$, then $Q^*(f) \geq \Omega(\log(\sigma_1))$.
If $\sigma_1 \leq \sqrt{\sigma_0}$, then $Q^*(f) \geq \Omega(\log(\sigma_1)/(\log\sqrt{\sigma_0} - \log\sigma_1 + 1))$.

PROOF. The proof is very similar to the proof of Theorem 21. □

PROOF OF THEOREM 23. Let B be a real matrix such that

1. $\gamma_2(B) = \gamma_2^{1+\delta}(A)$.
2. $1 \leq b_{ij} a_{ij} \leq 1 + \delta$ for all i, j.

Condition 2 implies that $\|A - B\|_\infty \leq \delta$, and hence $\|A - B\|_2 \leq \delta 2^n$.
By Parseval identity

$$\sqrt{\sum_{(z,z)\in E} \left(\hat{A}_{z,z} - \hat{B}_{z,z}\right)^2} \leq 2^{-n}\|A - B\|_2 \leq \delta.$$

By the triangle inequality and Cauchy-Schwartz

$$\begin{aligned}\sum_E |\hat{B}_{z,z}| &\geq \sum_E |\hat{A}_{z,z}| - \sum_E |\hat{A}_{z,z} - \hat{B}_{z,z}| \\ &\geq \sum_E |\hat{A}_{z,z}| - \sqrt{|E| \cdot \sum_E \left(\hat{A}_{z,z} - \hat{B}_{z,z}\right)^2} \\ &\geq \sigma_1 - \sqrt{\sigma_0} \cdot \delta.\end{aligned}$$

By Lemma 6 it is possible to express $B = \sum_i w_i x_i y_i^t$, where $w_1, \ldots, w_s$ are positive reals with $\sum w_i \leq K_G \delta$ and $x_1, \ldots, x_s, y_1, \ldots, y_s$ are sign vectors. Using Observation 4 and the linearity of the Fourier transform, we obtain

$$\begin{aligned}\sum_E |\hat{B}_{z,z}| &= \sum_E \sum_i |w_i \hat{x}_{i,z} \hat{y}_{i,z}| \\ &= \sum_i w_i \sum_E |\hat{x}_{i,z} \hat{y}_{i,z}| \leq \sum_i w_i,\end{aligned}$$

where the inequality holds since $\hat{x}, \hat{y}$ are unit vectors. We conclude that

$$\sigma_1 - \sqrt{\sigma_0} \cdot \delta \leq \sum_E |\hat{B}_{z,z}| \leq \sum_i w_i \leq K_G \gamma_2^{1+\delta}(A),$$

as claimed. □

4.3.2 Using a single Fourier coefficient

For every function $f : \mathbb{Z}_2^n \to \{\pm 1\}$, we denote by $\Lambda_f = (\lambda_{xy})$ the $2^n \times 2^n$ matrix with $\lambda_{xy} = f(x \wedge y)$. It was proved by Klauck [6] that

THEOREM 25. *For every function $f : \mathbb{Z}_2^n \to \{\pm 1\}$ and all $z \in \mathbb{Z}_2^n$*

$$Q(\Lambda_f) \geq \Omega\left(\frac{|z|}{1 - \log|\hat{f}_z|}\right).$$

(Here and below $|z|$ stands for the Hamming weight of z). He also asked whether the same lower bound holds when entanglement is allowed. We show that this is indeed the case, namely:

THEOREM 26. *For every function $f : \mathbb{Z}_2^n \to \{\pm 1\}$ and all $z \in \mathbb{Z}_2^n$*

$$Q^*(\Lambda_f) \geq \Omega\left(\frac{|z|}{1 - \log|\hat{f}_z|}\right).$$

The main part of the proof consists of showing:

THEOREM 27. *For every function $f : \mathbb{Z}_2^n \to \{\pm 1\}$ and all $z \in \mathbb{Z}_2^n$*

$$\gamma_2^{1+|\hat{f}_z|/2}(\Lambda_f) \geq \Omega\left(2^{|z|/4}|\hat{f}_z|\right).$$

PROOF OF THEOREM 27. We assume w.l.o.g. that $\hat{f}_z \geq 0$, to simplify the notations.

As stated in Theorem 12, for every sign matrix A,

$$\gamma_2^\alpha(A) = \max \quad \langle A, B\rangle - (\alpha - 1)\sum_{xy: a_{xy} \neq sign(b_{xy})} |b_{xy}| \quad \text{s.t.} \quad \gamma_2^*(B) \leq 1 \quad .$$

The proof proceeds by selecting for each $z \in \mathbb{Z}_2^n$ a matrix $B = B_z$ to yield the desired lower bound. We first describe this choice of B, and then apply it toward the lower bound.

Let $P = P_n$ be the $2^n \times 2^n$ matrix, with rows and columns indexed by vectors in $\{0,1\}^n$, where the x, y entry is

$$(\frac{1}{\sqrt{2}})^{|x|}(1 - \frac{1}{\sqrt{2}})^{n-|x|}(\frac{1}{\sqrt{2}})^{|y|}(1 - \frac{1}{\sqrt{2}})^{n-|y|}.$$

For what follows it is useful to observe that P induces a product probability distribution on $2^{[n]} \times 2^{[n]}$, each probability distribution being itself a bitwise product distribution. It has the property that for every $w \in \{0,1\}^n$, the event $\{(x,y) \in 2^{[n]} \times 2^{[n]}, \text{ s.t. } x \wedge y = w\}$ has probability 2^{-n}. For $z \in \mathbb{Z}_2^n$ we choose $B_z = P_n \circ \Lambda_{\chi_z}$. It is useful to observe

that $\Lambda_{\chi_z} = H_{|z|} \otimes J_{n-|z|}$, where H_t is the $2^t \times 2^t$ Sylvester-Hadamard matrix, and J_t is the $2^t \times 2^t$ matrix whose entries are all 1.

To apply Theorem 12 we need to compute (or estimate) $\gamma_2^*(B_z)$, and $\langle A, B_z \rangle$. Indeed,

1. For every $z \in \mathbb{Z}_2^n$, $\langle B_z, \Lambda_f \rangle = \hat{f}_z$.
2. There is a constant $c > 0$ such that for every $z \in \mathbb{Z}_2^n$
$$\gamma_2^*(B_z) \leq c 2^{-|z|/4}.$$

For the first equality, observe that

$$\begin{aligned} \langle B_z, \Lambda_f \rangle &= \sum_{x,y} P(x \wedge y) f(x \wedge y) \chi_z(x \wedge y) \\ &= \frac{1}{2^n} \sum_w f(w) \chi_z(w) = \hat{f}_z \end{aligned}$$

As for the second inequality - It follows from a similar inequality from [6] on the $\|\cdot\|_{\infty \to 1}$ norm. The additional step is provided by Inequality (4). It is left to compute the result of applying B_z. Let $B_z = (b_{xy})$ then $\gamma_2^{1+\hat{f}_z/2}(\Lambda_f)$ is at most

$$c^{-1} 2^{|z|/4} \left(\langle \Lambda_f, B_z \rangle - \frac{\hat{f}_z}{2} \sum_{xy: \lambda_{xy} \neq sign(b_{xy})} |b_{xy}| \right),$$

consequently

$$\begin{aligned} \gamma_2^{1+\hat{f}_z/2}(\Lambda_f) &\geq c^{-1} 2^{|z|/4} \left(\hat{f}_z - \frac{\hat{f}_z}{2} \|B_z\|_1 \right) \\ &= c^{-1} 2^{|z|/4} \left(\hat{f}_z - \frac{\hat{f}_z}{2} \right) \\ &= c^{-1} 2^{|z|/4} \hat{f}_z / 2. \end{aligned}$$

The third equality follows since $B_z = P_n \circ \Lambda_{\chi_z}$ is obtained by signing (via Λ_{χ_z} - a sign matrix) the terms of a probability distribution - the entries of P. □

Proof of Theorem 26. We use Theorem 27. By taking the logarithm in Theorem 27, we obtain

$$\log(\gamma_2^{1+|\hat{f}_z|/2}(\Lambda_f)) \geq |z|/4 + \log|\hat{f}_z| - O(1).$$

By Theorem 11

$$Q_\epsilon^*(\Lambda_f) \geq \log \gamma_2^{\alpha_\epsilon}(\Lambda_f) - \log \alpha_\epsilon - 2,$$

for any $\epsilon > 0$ where $\alpha_\epsilon = \frac{1}{1-2\epsilon}$.

We apply this with $\epsilon = \frac{|\hat{f}_z|}{4+2|\hat{f}_z|}$ (whence $\alpha_\epsilon = 1 + |\hat{f}_z|/2$). The two inequalities combined yield

$$Q_\epsilon^*(\Lambda_f) \geq |z|/4 - \log|\hat{f}_z| - \log \alpha_\epsilon - O(1).$$

As already mentioned, by a standard amplification argument (e.g. [9]),

$$Q^*(\Lambda_f) \geq \Omega\left(\frac{Q_\epsilon^*(\Lambda_f)}{\log \epsilon^{-1}} \right).$$

This yields

$$Q^*(\Lambda_f) \geq \Omega\left(\frac{|z|/4 + \log|\hat{f}_z| - \log \alpha_\epsilon - O(1)}{\log \epsilon^{-1}} \right).$$

Theorem 26 follows when we notice that $\epsilon = \Theta(|\hat{f}_z|)$ and $-\log \alpha_\epsilon = \Theta(1)$. □

4.4 Entropy

The *entropy* of a probability vector p is denoted $H(p) = -\sum_i p_i \log p_i$. Let B be an $n \times n$ real matrix, recall (e.g., [2, p. 7]) that $\sum_i s_i(B)^2 = \|B\|_2^2$. Thus, if we denote $\hat{s}_i(B) = \frac{s_i(B)}{\|B\|_2}$ then the vector $\hat{s}(B)^2 = (\hat{s}_1(B)^2, \ldots, \hat{s}_n(B)^2)$ is a probability vector. Klauck [6] proved

Theorem 28. *For every $n \times n$ sign matrix A*

$$Q(A) \geq \Omega\left(\frac{H(\hat{s}(A)^2)}{\log \log n} \right).$$

We generalize Klauck's result

Theorem 29. *For every sign matrix A and $\delta \leq 1/6$*

$$\log\left(1 + \gamma_2^{1+\delta}(A)\right) \geq \frac{1}{2} H(\hat{s}(A)^2) - \frac{3}{2}\delta \cdot \log n.$$

By optimizing the choice of δ in Theorem 29, Theorem 11 yields the following theorem (see the proof of Theorem 21, which is very similar, for details)

Theorem 30. *For every $n \times n$ sign matrix A*

$$Q^*(A) \geq \Omega\left(\frac{H(\hat{s}(A)^2)}{\log \frac{\log n}{H(\hat{s}(A)^2)} + 1} \right).$$

Proof of Theorem 29. We use the following simple properties of entropy:

Lemma 31 ([6]). *Let p and q be probability vectors of dimension n, then*

1. *If $\|p-q\|_1 \leq 1/2$ then $|H(p) - H(q)| \leq \|p-q\|_1 \cdot \log n - O(1)$.*
2. *$\|p-q\|_1 \leq 3\|p^{1/2} - q^{1/2}\|_2$. Here $p^{1/2} = (\sqrt{p_1}, \ldots, \sqrt{p_n})$.*
3. *$H(p) \leq 2 \log\left(1 + \|p^{1/2}\|_1\right)$.*

For $\delta \leq 1/6$, let B be a real matrix satisfying $\gamma_2(B) = \gamma_2^{1+\delta}(A)$ and $1 \leq a_{ij} b_{ij} \leq 1 + \delta$. By property (3) in Lemma 31,

$$\begin{aligned} H(\hat{s}(B)^2) &\leq 2 \log\left(1 + \frac{\|B\|_{tr}}{\|B\|_2}\right) \\ &\leq 2 \log\left(1 + \frac{\|B\|_{tr}}{n}\right) \\ &\leq 2 \log\left(1 + \gamma_2(B)\right) \\ &= 2 \log\left(1 + \gamma_2^{1+\delta}(A)\right). \end{aligned} \tag{15}$$

By the second property

$$\begin{aligned} \|\hat{s}(A)^2 - \hat{s}(B)^2\|_1 &\leq 3\|\hat{s}(A) - \hat{s}(B)\|_2 \\ &= 3\|s(A/\|A\|_2) - s(B/|B\|_2)\|_2 \\ &\leq 3\|A/\|A\|_2 - B/|B\|_2\|_2 \\ &\leq \frac{3}{\|A\|_2}\|A - B\|_2 \\ &\leq \frac{3}{n}\delta \cdot n \\ &= 3\delta. \end{aligned}$$

For the second inequality see Theorem VI.4.1 and Exercise II.1.15 in [2]. The third inequality follows from the simple fact that $\|\frac{y}{\|y\|_2} - \frac{x}{\|x\|_2}\|_2 \leq \frac{\|y-x\|_2}{\|x\|_2}$ for every two vectors with $\|y\|_2 \geq \|x\|_2$ (Here $x = A$ and $y = B$). Notice

that $\|\hat{s}(A)^2 - \hat{s}(B)^2\|_1 \leq 3\delta \leq 1/2$, the conditions of the first property in Lemma 31 are therefore satisfied, and we have

$$\begin{aligned} H(\hat{s}(B)^2) &\geq H(\hat{s}(A)^2) - \|\hat{s}(A)^2 - \hat{s}(B)^2\|_1 \cdot \log n - O(1) \\ &\geq H(\hat{s}(A)^2) - 3\delta \cdot \log n - O(1). \end{aligned}$$

Combining this with (15), the bound in the theorem is proved. □

5. EXAMPLES

This section contains examples that exhibit gaps between different complexity measures considered in this paper.

5.1 Disjointness matrix

Many of the concrete examples analyzed in the literature on communication complexity are symmetric functions. In particular - the disjointness function. Let $D_k = (d_{xy})$ be a $2^k \times 2^k$ matrix with rows and columns indexed by the subsets of $[k]$, where

$$d_{xy} = \begin{cases} 1 & \text{if } x \cap y \neq \emptyset \\ -1 & \text{if } x \cap y = \emptyset \end{cases} \tag{16}$$

There is a rich literature concerning the communication complexity of this function. It is particularly interesting in the context of the present paper because the various proof techniques mentioned here vary significantly in the bounds they yield for the disjointness function. We now recall some of the key parameters of the disjointness matrix, and see what they imply for the complexity measures at hand. The relevant references or proofs are then provided.

1. $disc(D_k)^{-1} \leq O(\gamma_2^\infty(D_k)) \leq O(k)$.
2. For $\alpha = 3/2$, $2^{\tilde{O}(\sqrt{k})} \geq Q^*(D_k) \geq \gamma_2^\alpha(D_k) \geq \|D_k\|_{tr}^\alpha / 2^k \geq 2^{\tilde{\Omega}(\sqrt{k})}$. (Here and below tildes indicate missing log factors).
3. $o(2^{k/2}) \geq \gamma_2(D_k) \geq \|D_k\|_{tr}/2^k \geq \left(\frac{\sqrt{5}}{2}\right)^k - 1$.

It follows from properties (1-3) that $\gamma_2^\alpha(D_k)$ decreases very rapidly as α grows. In particular, this is an example where γ_2 is much larger than γ_2^α even for small α, and there is an exponential gap between $\gamma_2^{3/2}$ and γ_2^∞ (equivalently, the inverse of discrepancy). It is interesting to better understand the behavior of γ_2 as a function of α. Furthermore, the disjointness matrix is also an example where the bound via the trace norm of Theorem 18 is exponentially better than the discrepancy bound.

We turn to discuss the first item. The discrepancy of D_k can be estimated by a simple explicit construction. Let H_k be the $k \times 2^k$ $(0,1)$-matrix with no repeated columns, and $B = 2(H_k^t H_k) - J$. Namely $b_{xy} = 2|x \cap y| - 1$, whence $b_{xy} d_{xy} \geq 1$ for all x, y. Consequently,

$$\gamma_2^\infty(D_k) \leq \gamma_2(B) \leq 2k + 1.$$

(For the last calculation use the fact that γ_2 is a norm and that $\gamma_2(J) = 1$.)
It follows that

$$disc(D_k)^{-1} \leq O(\gamma_2^\infty(D_k)) \leq O(k).$$

On the other hand it follows from [16] that for $\alpha = 3/2$,

$$2^{\tilde{O}(\sqrt{k})} \geq Q^*(D_k) \geq \|D_k\|_{tr}^\alpha / 2^k \geq 2^{\tilde{\Omega}(\sqrt{k})}.$$

Combining this with Theorem 11 and the discussion in Section 4.2 we get the statement of (2) ($\gamma_2^\alpha(D_k)$ falls between $Q^*(D_k)$ and $\|D_k\|_{tr}^\alpha / 2^k$).

To estimate the trace norm of D_k and $\gamma_2(D_k)$ we introduce the matrix $E_k = \frac{1}{2}(D_k + J)$. We estimate the trace norm of E_k, and use the fact that $|\ \|D_k\|_{tr} - \|E_k\|_{tr}\ | \leq 2^k$. Observe that $E_k = E_1^{\otimes k}$, and that the singular values of E_1 are $\frac{\sqrt{5}\pm 1}{2}$. The 2^k singular values of E_k consist of all the numbers expressible as the product of k terms, each of which is either $\frac{1+\sqrt{5}}{2}$ or $\frac{\sqrt{5}-1}{2}$. Therefore, by the binomial identity $\|E_k\|_{tr} = \|E_1\|_{tr}^k = (\sqrt{5})^k$, and

$$\gamma_2(D_k) \geq \|D_k\|_{tr}/2^k \geq \left(\frac{\sqrt{5}}{2}\right)^k - 1.$$

Finally, it follows from Claim 13 and property (2) that $\gamma_2(D_k) \leq o(2^{k/2})$, since if it were the case that $\gamma_2(D_k) = \Omega(2^{k/2})$, then by Claim 13 also $\gamma_2^{3/2}(D_k) = \Omega(2^{k/2})$ contradicting property (2).

5.2 γ_2 vs. the trace norm

It is shown in [10] that $\gamma_2^\infty(H) = \sqrt{m}$ for an $m \times m$ Hadamard matrix H. For $n = \Theta(m^{3/2})$ let Z be an $n \times n$ matrix with H as a principal minor and all other entries equal to 1. It is not hard to check that for every $\alpha \geq 1$

$$1 \geq \|Z\|_{tr}/n \geq \|Z\|_{tr}^\alpha / n,$$

while

$$\gamma_2^\alpha(Z) \geq \gamma_2^\infty(Z) \geq O(n^{1/3}).$$

So the inverse of discrepancy can be much larger than $\|\cdot\|_{tr}^{\alpha_\epsilon}$. In such cases Theorem 11 gives a bound that is significantly better than Theorem 18. Also, combining this with the example in Section 5.1 we see that there is no general inequality between the inverse of discrepancy and $\|\cdot\|_{tr}^{\alpha_\epsilon}$ and either one can be significantly larger than the other.

6. DISCUSSION AND OPEN PROBLEMS

The results of this paper show that deep properties of communication protocols can be investigated using factorization norms. However, many questions in this area remain open. Already in Section 3.4 we asked:

QUESTION 32. *Is it true that for every sign matrix A there holds $R_{1/3}(A) \geq \Omega(\log \gamma_2(A))$?*

Another question concerns upper bounds on communication complexity in terms of factorization norms. Claim 15 bounds the randomized communication complexity from above by a power of γ_2^∞. The bound is tight, as stated, but it is conceivable that much tighter upper bounds hold, if we consider γ_2^α instead. Perhaps even a power of $\log(\gamma_2^\alpha)$ suffices? This raises the following problem

PROBLEM 33. *Find the best upper bound on randomized communication complexity in terms of γ_2^α. In particular, is there a constant k such that $R(A) \leq (\log(\gamma_2(A)))^k$ for every sign matrix A?*

In view of Proposition 3, this problem is analogous to the *log rank conjecture* [13, 12], which asks whether

$$CC(A) \leq (\log rank(A))^k,$$

for some constant k and for every sign matrix A. Here CC stands for deterministic communication complexity. [3] Lovász and Saks [12], proved the log rank conjecture in some special cases. On the other hand, an example due to Nisan and Wigderson [13] shows that if this conjecture is true, then necessarily $k \geq \log_2 3$. We note that the same example implies that in the latter part of Problem 33 k must be at least $\log_2 3$ as well.

Problem 33 raises the intriguing possibility that randomized communication complexity and γ_2 are closely related. An affirmative answer would be rather surprising, in view of the fact that the two notions seem a priori unrelated. A resolution of this question would presumably require some new and interesting ideas. It is also interesting to note the relation between this question and work by Grolmusz [4].

Our final question is this:

PROBLEM 34. *Fix a sign matrix A and consider $\gamma_2^\alpha(A)$ as a function of α. What can be said about the behavior of such functions? Specifically what are the relationships between $\gamma_2 = \gamma_2^1$ and γ_2^∞ ?*

This function of α is, of course, decreasing and convex but very little is known in general, and even very special cases, such as $A = D_k$, seem interesting and challenging.

Some information about the possible gap between $\gamma_2 = \gamma_2^1$ and γ_2^∞ can be found in [10] and the present paper say a little more about this question. Namely, combining the results of Theorem 9, Claim 15 and Lemma 14 we conclude that if A is an $n \times n$ sign matrix with $\gamma_2(A) \geq \Omega(\sqrt{n})$ then $\gamma_2^\infty(A) \geq \Omega(\sqrt{\log n})$.

Acknowledgments

We thank Julia Kempe and Ronald de Wolf for helpful comments, and Gideon Schechtman for many fruitful discussions.

[3] As mentioned, $\log(rank(A)) \leq CC(A)$ for every sign matrix A.

7. REFERENCES

[1] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity. In *Proceedings of the 27th IEEE FOCS*, pages 337–347, 1986.

[2] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.

[3] D. Gavinsky, J. Kempe, and R. de Wolf. Strength and weaknesses of quantum fingerprinting, 2006. accepted to CCC"06.

[4] V. Grolmusz. Harmonic analysis, real approximation, and the communication complexity of boolean functions. *Algorithmica*, 23(4):341–353, 1999.

[5] G. J. O. Jameson. *Summing and nuclear norms in Banach space theory*. London mathematical society student texts. Cambridge university press, 1987.

[6] H. Klauck. Lower bounds for quantum communication complexity. In *Proceedings of the 42nd IEEE FOCS*, pages 288–297, 2001.

[7] I. Kremer. Quantum communication. In *Master's thesis*. Hebrew University of Jerusalem, 1995.

[8] I. Kremer, N. Nisan, and D. Ron. On randomized one-round communication complexity. In *Proceedings of the 35th IEEE FOCS*, 1994.

[9] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambride University Press, 1997.

[10] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman. Complexity measures of sign matrices. *Combinatorica*, to appear. Available at `http://www.cs.huji.ac.il/~nati/PAPERS/complexity_matrices.ps.gz`.

[11] N. Linial and A. Shraibman. Learning complexity vs. communication complexity. *Manuscript*, 2006.

[12] L. Lovasz and M. Saks. Latices, Mobius functions, and communication complexity. In *Proceedings of the 29th IEEE FOCS*, pages 81–90, 1988.

[13] N. Nisan and A. Wigderson. On rank vs. communication complexity. In *Proceedings of the 35th IEEE FOCS*, pages 831–836, 1994.

[14] G. Pisier. *Factorization of linear operators and geometry of Banach spaces*, volume 60 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1986.

[15] R. Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, 5(3/4):205–221, 1995.

[16] A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Science, Mathematics*, 67:145–159, 2002.

[17] A. Yao. Lower bounds by probabilistic arguments. In *Proceedings of the 15th ACM STOC*, pages 420–428, 1983.

[18] A. Yao. Quantum circuit complexity. In *Proceedings of the 34th IEEE FOCS*, pages 352–361, 1993.