Separation of the Factorization Norm and Randomized Communication Complexity

Tsun-Ming Cheung
School of Computer Science, McGill University, Montreal, Canada

Hamed Hatami
School of Computer Science, McGill University, Montreal, Canada

Kaave Hosseini
Department of Computer Science, University of Rochester, NY, USA

Morgan Shirley
Department of Computer Science, University of Toronto, Canada

Abstract

In an influential paper, Linial and Shraibman (STOC '07) introduced the factorization norm as a powerful tool for proving lower bounds against randomized and quantum communication complexities. They showed that the logarithm of the *approximate* γ_2-factorization norm is a lower bound for these parameters and asked whether a stronger lower bound that replaces approximate γ_2 norm with the γ_2 norm holds.

We answer the question of Linial and Shraibman in the negative by exhibiting a $2^n \times 2^n$ Boolean matrix with γ_2 norm $2^{\Omega(n)}$ and randomized communication complexity $O(\log n)$.

As a corollary, we recover the recent result of Chattopadhyay, Lovett, and Vinyals (CCC '19) that deterministic protocols with access to an Equality oracle are exponentially weaker than (one-sided error) randomized protocols. In fact, as a stronger consequence, our result implies an exponential separation between the power of unambiguous nondeterministic protocols with access to Equality oracle and (one-sided error) randomized protocols, which answers a question of Pitassi, Shirley, and Shraibman (ITSC '23).

Our result also implies a conjecture of Sherif (Ph.D. thesis) that the γ_2 norm of the Integer Inner Product function (`IIP`) in dimension 3 or higher is exponential in its input size.

2012 ACM Subject Classification Theory of computation → Communication complexity

Keywords and phrases Factorization norms, randomized communication complexity

Digital Object Identifier 10.4230/LIPIcs.CCC.2023.1

Funding *Hamed Hatami*: Supported by an NSERC grant.
Morgan Shirley: Supported by an NSERC grant.

1 Introduction

The γ_2-factorization norm is an important notion of matrix complexity that was initially developed in Banach Space theory. In an influential paper, Linial and Shraibman [12] introduced this norm to communication complexity. Subsequently, the factorization norm and its approximate version found numerous applications in communication complexity and other adjacent areas such as discrepancy theory [13] and differential privacy [14, 3, 7].

▶ **Definition 1** (γ_2-factorization norm). *The* γ_2 norm *of a real matrix* A *is*

$$\|A\|_{\gamma_2} := \min_{X,Y:A=XY} \|X\|_{\text{row}} \|Y\|_{\text{col}},$$

where $\|X\|_{\text{row}}$ *and* $\|Y\|_{\text{col}}$ *denote the largest* ℓ_2*-norm of a row in* X *and the largest* ℓ_2 *norm of a column in* Y*, respectively.*


▶ **Definition 2** (Approximate γ_2 norm). *The* approximate γ_2 norm *of $A \in \mathbb{R}^{k\times\ell}$ with error ϵ, denoted by $\widetilde{\gamma}_2^{\epsilon}(A)$, is the minimum $\|B\|_{\gamma_2}$ over all matrices $B \in \mathbb{R}^{k\times\ell}$ with $\|A-B\|_\infty \leq \epsilon$.*

We use the notation $\widetilde{\gamma}_2^{\epsilon}(\cdot)$ to emphasize that unlike the γ_2 norm $\|\cdot\|_{\gamma_2}$, the approximate γ_2 norm is not a norm. The choice of the error parameter ϵ is mostly unimportant in the context of communication complexity. Indeed, a constant-factor reduction in the error parameter increases $\log \widetilde{\gamma}_2^{\epsilon}(A)$ by a constant factor [1, Lemma 21]. Therefore, we use the standard choice of $\epsilon = 1/3$ and write $\widetilde{\gamma}_2$ for $\widetilde{\gamma}_2^{1/3}$. Both of the quantities $\widetilde{\gamma}_2$ and γ_2 are polynomial-time computable using semi-definite programming [12].

Linial and Shraibman [12] showed that $\log \widetilde{\gamma}_2(A)$ provides a lower bound on the public-coin randomized communication complexity $\mathrm{R}(A)$ and the quantum communication complexity with shared entanglement $\mathrm{Q}^*(A)$:

$$\log \widetilde{\gamma}_2(A) \lesssim \mathrm{Q}^*(A) \leq \mathrm{R}(A). \tag{1}$$

These lower bounds subsume the most well-known lower bounds on randomized and quantum communication complexity, such as discrepancy, approximate trace norm [17], and entropy of singular values [9].

Linial and Shraibman [12] state that "they cannot rule out the intriguing possibility that the first inequality in Equation (1) is a tip of something bigger and randomized communication complexity and the quantum communication complexity with shared entanglement are in fact polynomially equivalent to $\log \|A\|_{\gamma_2}$."

▶ **Question 1** ([12]). *Is $\log \|A\|_{\gamma_2} \leq \widetilde{O}(\mathrm{R}(A))$ for every a Boolean matrix $A : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$?*

Here, the notation $\widetilde{O}(\cdot)$ hides a factor of $\mathrm{polylog}(n)$, which is common in communication complexity since the communication cost of $\mathrm{polylog}(n)$ is considered efficient.

Another motivation for Question 1 comes from the following observation. It is well-known that the Equality function $\textsc{eq} : \{0,1\}^n \times \{0,1\}^n$ has $\mathrm{R}(\textsc{eq}) = O(1)$ (see e.g. [10]) but its rank over the reals is 2^n, and therefore EQ witnesses the strongest possible separation ($O(1)$ versus 2^n) between R and rank. On the other hand, as mentioned before, the γ_2 norm can be viewed as a smooth analogue of rank. However, the γ_2 norm of the Equality function is 1, and therefore, one naturally wonders whether there is a strong separation between $\mathrm{R}(\cdot)$ and the γ_2 norm.

The purpose of the present paper is to give a strong negative answer to Question 1. In fact, we work with a stronger parameter of $\mathrm{R}_0^1(A)$ instead of $\mathrm{R}(A)$. This parameter is the minimum cost of a *one-sided* public-coin randomized protocol. The protocol is not allowed to have any error on 1 entries of A, but on the 0 entries, it can have a probability of error as big as 1/3.

1.1 Main Result

Our main result establishes a strong separation between the γ_2 norm and R_0^1.

▶ **Theorem 3** (Main Theorem). *There is a Boolean matrix $M : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with $\|M\|_{\gamma_2} \geq 2^{n/32}$ and $\mathrm{R}_0^1(M) \leq O(\log n)$.*

The construction in Theorem 3 is based on the *point-line incidence matrix* over the integers. For integers $1 \leq q \leq p$, let $\mathtt{PL}$ be the $qp \times qp$ Boolean matrix whose rows and columns are indexed by the elements of $[q] \times \{0,\ldots,p-1\}$ and its entries are given as $\mathtt{PL}[(x,x'),(y,y')] = 1$ iff $xy + x' = y'$. We also define a variant of $\mathtt{PL}$ over $\mathbb{Z}_p$ to simplify the analysis. The matrix $\mathtt{PL}_{\mathbb{Z}_p}$ is the $qp \times qp$ Boolean matrix whose rows and columns are indexed by $[q] \times \mathbb{Z}_p$ and its entries are given as $\mathtt{PL}_{\mathbb{Z}_p}[(x,x'),(y,y')] = 1$ iff $xy + x' \equiv y' \mod p$.

Recall that the trace norm of a matrix is the sum of its singular values (see Section 2.1).Theorem 3 is immediate from the following theorem, which is our main technical contribution.

▶ **Theorem 4** (Technical Statement of the Main Theorem). *Let p be a prime.*

(i) *For $1 \le q \le \sqrt{p}$, we have*

$$\|\mathtt{PL}_{\mathbb{Z}_p}\|_{\mathrm{Tr}} = \Omega(pq^{9/8}) \quad \textit{and} \quad \|\mathtt{PL}_{\mathbb{Z}_p}\|_{\gamma_2} = \Omega(q^{1/8}) \quad \textit{and} \quad \mathrm{R}_0^1(\mathtt{PL}_{\mathbb{Z}_p}) = O(\log\log p).$$

(ii) *For $1 \le q \le p^{1/3}$, we have*

$$\|\mathtt{PL}\|_{\mathrm{Tr}} = \Omega(pq^{9/8}) \quad \textit{and} \quad \|\mathtt{PL}\|_{\gamma_2} = \Omega(q^{1/8}) \quad \textit{and} \quad \mathrm{R}_0^1(\mathtt{PL}) = O(\log\log p).$$

▶ Remark 5. The condition $1 \le q \le p^{1/3}$ in (ii) allows us to deduce (ii) from (i) since $\|\mathtt{PL}_{\mathbb{Z}_p} - \mathtt{PL}\|_{\mathrm{Tr}} = o(pq^{9/8})$ in this range (see Lemma 15). On the other hand, the condition $1 \le q \le \sqrt{p}$ in (i) is to guarantee $\mathrm{R}_0^1(\mathtt{PL}_{\mathbb{Z}_p}) = O(\log\log p)$. Indeed, unlike $\mathtt{PL}$, whose randomized communication complexity is always small, the randomized communication complexity of $\mathtt{PL}_{\mathbb{Z}_p}$ is large when q is close to p. For example, for $q = p$, this follows from the fact that all nontrivial eigenvalues of $\mathtt{PL}_{\mathbb{Z}_p}$ are at most $\sqrt{3p}$ [19].

1.2 Consequences of the Main Theorem

As an immediate consequence, combining Theorem 3 with Equation (1) implies an exponential separation between $\widetilde{\gamma}_2(\cdot)$ and $\|\cdot\|_{\gamma_2}$. This corollary answers a question of Pitassi, Shirley, and Shraibman [16, Open Question 3].

▶ **Corollary 6.** *There is a Boolean matrix $M : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with $\|M\|_{\gamma_2} \ge 2^{n/32}$ and $\widetilde{\gamma}_2(M) \le O(\mathrm{poly}(n))$.*

Another corollary of Theorem 3 concerns the deterministic communication complexity with oracle access to the Equality function. We formally define this model in Section 2.2 and denote the corresponding complexity measure by $\mathrm{D}^{\mathrm{EQ}}(\cdot)$. The equality function, which corresponds to the identity matrix, is the standard example of a problem with $O(1)$ randomized communication complexity but large deterministic communication complexity. This fact makes $\mathrm{D}^{\mathrm{EQ}}(\cdot)$ an interesting complexity measure between randomized and deterministic communication complexities.

$$\log \widetilde{\gamma}_2(A) \lesssim \mathrm{Q}^*(A) \le \mathrm{R}(A) \lesssim \mathrm{D}^{\mathrm{EQ}}(A) \le \mathrm{D}(A). \tag{2}$$

Since the γ_2 norm of the identity matrix is 1, it is not hard to see that [5, Proposition 3.1]

$$\frac{1}{2}\log \|A\|_{\gamma_2} \le \mathrm{D}^{\mathrm{EQ}}(A). \tag{3}$$

In light of Equation (3), Theorem 3 implies the following.

▶ **Corollary 7.** *There is a Boolean matrix $M : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with $\mathrm{R}_0^1(M) \le O(\log n)$ and $\mathrm{D}^{\mathrm{EQ}}(M) = \Omega(n)$.*

The above corollary recovers the result of Chattopadhyay, Lovett, and Vinyals [2] separating R and D^{EQ}. In fact, we obtain an exponential lower bound on a model stronger than D^{EQ}. In complexity theory, *unambiguous nondeterminism* is similar to nondeterminism but with the extra requirement that for every input, there is at most one accepting computational path. Therefore, the power of unambiguous nondeterminism lies between determinism and nondeterminism. For a Boolean matrix M, the *unambiguous nondeterministic communication complexity of M with access to an equality oracle* is denoted by $\mathrm{UP}^{\mathrm{EQ}}$ (see Section 2.2). It is immediate that $\mathrm{UP}^{\mathrm{EQ}}(\cdot) \le \mathrm{D}^{\mathrm{EQ}}(\cdot)$. Theorem 3 implies the following corollary, answering a question of Pitassi, Shirley, and Shraibman [16, Open Question 2].

▶ **Corollary 8.** *There is a Boolean matrix* $M : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ *with* $\mathrm{R}_0^1(M) \leq O(\log n)$ *and* $\mathrm{UP}^{EQ}(M) = \Omega(n)$.

The matrix PL that we consider in Theorem 4 is essentially a submatrix of the Integer Inner Product matrix (IIP) used in the work of Chattopadhyay et al. [2]; however, the proof technique here is entirely different.

▶ **Definition 9.** *Let* $t \in \mathbb{N}$ *be a fixed constant. For a positive integer* $m = 2^n$*, the Integer Inner Product function* $\mathtt{IIP}_t : \{-m, \ldots, m\}^t \times \{-m, \ldots, m\}^t \to \{0,1\}$ *is defined as*

$$\mathtt{IIP}_t[(x_1, \ldots, x_t), (y_1, \ldots, y_t)] = 1 \text{ iff } x_1y_1 + \ldots + x_ty_t = 0.$$

Since t is a fixed constant, the input size of $\mathtt{IIP}_t$ is $\Theta(n)$-bits as a communication problem. Chattopadhyay, Lovett, and Vinyals proved that $\mathrm{R}_0^1(\mathtt{IIP}_t) = O(\log n)$, and $\mathrm{D}^{\mathrm{EQ}}(\mathtt{IIP}_t) = \Omega(n)$ for $t \geq 6$.

Later, Sherif [18] conjectured $\|\mathtt{IIP}_t\|_{\gamma_2} = 2^{\Omega(n)}$ for $t \geq 6$. Since the matrix PL is a submatrix of $\mathtt{IIP}_3$, as a corollary of Theorem 4, we answer Sherif's question in the affirmative.

▶ **Corollary 10.** *For* $t \geq 3$*,*

$$\|\mathtt{IIP}_t\|_{\gamma_2} = 2^{\Omega(n)}.$$

Proof. Choose n such that $2^{n-1} \leq p \leq 2^n$ and $q = \lceil p^{1/3} \rceil$. From Theorem 4, we obtain PL as a submatrix of $\mathtt{IIP}_3$ with $m = 2^n$ such that $\|\mathtt{PL}\|_{\gamma_2} = \Omega(2^{n/32})$. Since the γ_2 norm cannot increase when restricting to a submatrix, we conclude that

$$\|\mathtt{IIP}_t\|_{\gamma_2} \geq \|\mathtt{IIP}_3\|_{\gamma_2} \geq \|\mathtt{PL}\|_{\gamma_2} = 2^{\Omega(n)}. \qquad \blacktriangleleft$$

▶ Remark 11. The condition $t \geq 3$ is necessary as $\|\mathtt{IIP}_2\|_{\gamma_2} = O(1)$. To prove the latter, we use Equation (3) and show $\mathrm{D}^{\mathrm{EQ}}(\mathtt{IIP}_2) = O(1)$. Note that if $x_1y_1 + x_2y_2 = 0$ and $y_1, x_2 \neq 0$, then $\frac{x_1}{x_2} = -\frac{y_2}{y_1}$. To check this equation, Alice and Bob can call the Equality oracle on rational inputs $\frac{x_1}{x_2}$ and $-\frac{y_2}{y_1}$.

1.3 Connections to Fourier Algebra Norm

The sum of the absolute values of the Fourier coefficients of a function $f : \mathbb{Z}_2^n \to \mathbb{R}$ is called the *algebra norm* of f:

$$\|f\|_A \coloneqq \|\widehat{f}\|_1 = \sum_{a \in \mathbb{Z}_2^n} |\widehat{f}(a)|.$$

For any error parameter $\epsilon \in (0, 1/2)$, the ϵ-*approximate algebra norm* of $f : \mathbb{Z}_2^n \to \{0,1\}$ is

$$\widetilde{A}_\epsilon(f) \coloneqq \inf\{\|g\|_A \ : \ \|f - g\|_\infty \leq \epsilon\}.$$

It is possible to use the XOR operation to lift these norms to the γ_2 norm and the approximate γ_2 norm [12]: for the matrix $F : \mathbb{Z}_2^n \times \mathbb{Z}_2^n \to \{0,1\}$ defined by $F(x,y) = f(x \oplus y)$, we have

$$\|f\|_A = \|F\|_{\gamma_2} \qquad \text{and} \qquad \widetilde{A}_\epsilon(f) = \widetilde{\gamma}_2^\epsilon(F).$$

The communication complexity measures of F are related to the parity query complexity measures of f. For example, we have

$$\mathrm{R}(F) \leq 2\,\mathrm{rdt}^{\oplus}(f),$$

where $\mathrm{rdt}^{\oplus}(f)$ denotes the randomized parity decision tree complexity of f (see [5]).

Therefore, the class of XOR-lifted Boolean functions provide a rich collection of matrices for which the questions about the factorization norm reduce to simpler questions about the Fourier algebra norm. In this setting, one can ask the analog of Question 1.

▶ **Question 2** (Open Question). *Is* $\log \|f\|_A = \widetilde{O}(\mathrm{rdt}^{\oplus}(f))$ *for every Boolean function* $f : \mathbb{Z}_2^n \to \{0,1\}$*?*

By the above discussion, if we find a counter-example f to Question 2, then $F(x,y) \coloneqq f(x \oplus y)$ would be a counter-example to Question 1. However, Question 2 remains open. Indeed, our counter-example to Question 1 is not an XOR-lift.

Finally, let us comment on the stronger versions of Question 1 and Question 2, where we do not tolerate a polylog(n) factor, i.e., replace $\widetilde{O}(\cdot)$ with $O(\cdot)$. Let $B(n,r) \subseteq \{0,1\}^n$ denote the Hamming ball of radius r around the origin, i.e.,

$$B(n,r) \coloneqq \left\{ x \in \{0,1\}^n \ : \ \sum_{i=1}^{n} x_i \le r \right\}.$$

Note that the lifted function $F_{n,r}(x,y) = \mathbf{1}_{B(n,r)}(x \oplus y)$ corresponds to the hamming distance problem, whose communication complexity is well-understood. We have [8]

$$\mathrm{rdt}^{\oplus}(\mathbf{1}_{B(n,r)}) \le O(r \log r) \qquad \text{and} \qquad \mathrm{R}(F_{n,r}) \le O(r \log r).$$

On the other hand, for $r \le n/2$, the following bounds are known [5, Lemma 2.15] about the Fourier algebra norm of $\mathbf{1}_{B(n,r)}$:

$$e^{-r} \sqrt{\sum_{i=0}^{r} \binom{n}{i}} \le \left\|\mathbf{1}_{B(n,r)}\right\|_A = \|F_{n,r}\|_{\gamma_2} \le \sqrt{\sum_{i=0}^{r} \binom{n}{i}}.$$

Therefore, in the context of Question 2 and Question 1, taking $r = O(1)$ provides examples of $f : \{0,1\}^n \to \{0,1\}$ and $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with

$$\mathrm{rdt}^{\oplus}(f) = O(1) \qquad \text{and} \qquad \log \|f\|_A = \Theta(\log n),$$

and

$$\mathrm{R}(F) = O(1) \qquad \text{and} \qquad \log \|F\|_{\gamma_2} = \Theta(\log n).$$

Paper Organization

In Section 2, we discuss the preliminaries of matrix norms, communication complexity, and Fourier analysis. We give a brief overview of the proof strategy in Section 3. We present the proof of Theorem 4 in Sections 4 and 5. Finally, we discuss several open problems in Section 6.

2 Notations and Preliminaries

For a positive integer k, we denote $[k] \coloneqq \{1, \dots, k\}$. We use the shorthand notations $a \equiv_p b$ to denote $a \equiv b \bmod p$. For a set S, we use the indicator function notation $\mathbf{1}_S$, which is evaluated to 1 on x if $x \in S$ and 0 otherwise. All the logarithms in this paper are in base 2.

We adopt the standard computer science asymptotic notations and use the tilde asymptotic notations to hide poly-logarithmic factors. We write $f \lesssim g$ to denote $f(n) = O(g(n))$.

For a vector $v \in \mathbb{C}^k$, we denote the ℓ_2-norm of v by $\|v\|_2 = \sqrt{\sum_i |v_i|^2}$. We denote the all-1 matrix by $\mathbf{J}$.

2.1 Matrix Norms

For a complex-valued matrix $A \in \mathbb{C}^{k\times\ell}$, we denote the singular values of A by

$$\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_{\min(k,\ell)}(A) \geq 0.$$

We primarily work with the matrix norm family of *Schatten norms*. For $p \in [1,\infty]$, the Schatten-p norm of a matrix is the ℓ_p norm of the vector of its singular values. The particular cases of $p = 1, 2, \infty$ are frequently used, and these norms are commonly known as *trace norm*, *Frobenius norm*, and *spectral norm* respectively:

$$\|A\|_{\mathrm{Tr}} = \|A\|_{S_1} = \sum_i \sigma_i$$

$$\|A\|_F = \|A\|_{S_2} = \sqrt{\sum_i \sigma_i^2} = \sqrt{\sum_{i,j} |A_{ij}|^2}$$

$$\|A\| = \|A\|_{S_\infty} = \sigma_1 = \max_{x\in\mathbb{C}^\ell : \|x\|_2=1} \|Ax\|_2 = \max_{\substack{u\in\mathbb{C}^k, v\in\mathbb{C}^\ell \\ \|u\|_2=\|v\|_2=1}} u^*Av$$

Viewing Schatten p-norm as the ℓ_p norm of the singular value vector, one can obtain several useful properties inherited from ℓ_p norms. One such property is the monotonicity of Schatten p-norm in p: $\|A\|_{S_p} \geq \|A\|_{S_q}$ for $1 \leq p < q \leq \infty$.

Similar to the case of ℓ_p norm, for $p, q \in [1,\infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, the dual norm of $\|\cdot\|_{S_p}$ is $\|\cdot\|_{S_q}$. With the inner product on the matrix space $\mathbb{C}^{k\times\ell}$ defined by $\langle A, B\rangle = \mathrm{Tr}(A^*B) = \sum_{ij} \overline{A_{ij}} B_{ij}$, the Schatten p-norm admits the following dual norm characterization:

$$\|A\|_{S_p} = \max_{\|B\|_{S_q}=1} |\langle A, B\rangle|.$$

For the particular case of $p = 1$, this yields

$$|\langle A, B\rangle| \leq \|A\|_{\mathrm{Tr}} \|B\|.$$

In particular, by setting $B = A$, we have

$$\|A\|_F^2 \leq \|A\|_{\mathrm{Tr}} \|A\|. \tag{4}$$

Next, we discuss a reformulation of the γ_2 norm in terms of the trace norm. As shown in [11], for $A \in \mathbb{R}^{k\times\ell}$, we have

$$\|A\|_{\gamma_2} = \max_{\substack{u\in\mathbb{R}^k, v\in\mathbb{R}^\ell \\ \|u\|_2=\|v\|_2=1}} \|A \circ uv^T\|_{\mathrm{Tr}}.$$

Here $\circ$ denotes the Hadamard (or entrywise) product of two matrices: for $B, C \in \mathbb{R}^{k\times\ell}$, their product $B \circ C$ is the $m \times n$ matrix defined by $[B \circ C]_{ij} = B_{ij}C_{ij}$ for all i, j. It follows from the trace norm formulation of the γ_2 norm that

$$\|A\|_{\gamma_2} \geq \frac{1}{\sqrt{k\ell}} \|A\|_{\mathrm{Tr}}. \tag{5}$$

2.2 Communication Complexity

In the standard communication model, there are two parties and problems are modelled by functions $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ on finite domains $\mathcal{X}, \mathcal{Y}$. The two parties receive $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, respectively, and they exchange messages to compute $f(x, y)$. We often interpret f as a Boolean matrix indexed by $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

For a given $\epsilon \in (0, 1/2)$, we denote by $\mathrm{R}_\epsilon(f)$, the randomized communication complexity of f in the public-coin model with two-sided error $\epsilon > 0$. The one-sided versions, $\mathrm{R}^1_\epsilon(f)$ and $\mathrm{R}^1_{0,\epsilon}(f)$, restrict the error to be one-sided: $\mathrm{R}^1_\epsilon(f)$ does not allow any error on the inputs in $f^{-1}(0)$. Similarly, $\mathrm{R}^1_{0,\epsilon}(f)$ does not allow any error on the inputs in $f^{-1}(1)$. We refer the reader to [10] for the formal definitions. We use the canonical choice of $\epsilon = 1/3$ and drop ϵ in the notations in such cases. This choice is without loss of generality since the probability of error can be reduced to any constant $\epsilon' > 0$ by repeating the protocol a constant number of times and outputting the majority.

As mentioned, approximate norms are useful tools for studying communication complexity. The following well-known inequalities [6, Proposition A.2] connect approximate γ_2 norm with randomized communication complexity.

$$\log \widetilde{\gamma}_2(A) \leq \mathrm{R}(A) \leq O(\widetilde{\gamma}_2(A)^2). \tag{6}$$

Next, we define the *deterministic communication complexity with access to an equality oracle*. In this model, a protocol computing a Boolean matrix $A_{\mathcal{X}\times\mathcal{Y}}$ corresponds to a binary tree. Each non-leaf node v in the tree is labelled with two functions $a_v : \mathcal{X} \to \mathcal{Z}$ and $b_v : \mathcal{Y} \to \mathcal{Z}$ for a finite set $\mathcal{Z}$. Such a node v corresponds to the query $\textsc{eq}(a_v(x), b_v(y))$, which returns 1 if $a_v(x) = b_v(y)$ and 0 otherwise. Every input (x, y) naturally corresponds to a path from the tree's root to a leaf, and the leaf must be labelled with the correct value $A(x, y)$. The cost of the protocol is the depth of the tree. The deterministic communication complexity of the matrix A with access to an equality oracle, denoted by $\mathrm{D}^{\textsc{eq}}(A)$, is the smallest depth of such a protocol for A.

Consider a node v in an equality-oracle deterministic communication protocol as described above. Note that the matrix $B_v(x, y) \coloneqq \textsc{eq}(a_v(x), b_v(y))$ consists of a collection of all-1 submatrices with rows and columns disjoint. Such matrices are dubbed *blocky matrices* by [5]. The answer to the query at the node v will inform the parties whether the input (x, y) belongs to the support of B_v or the support of $\mathbf{J} - B_v$.

Consider a leaf ℓ of the protocol tree where the protocol outputs 1, and let $v_1, \ldots, v_d = \ell$ be the set of the nodes on the corresponding path from the root. The inputs that lead the protocol to reach ℓ are the 1 entries of the matrix $M_\ell \coloneqq C_{v_1} \circ \ldots \circ C_{v_{d-1}}$ with $C_{v_i} = B_{v_i}$ or $C_{v_i} = \mathbf{J} - B_{v_i}$ according to the outcome of the query at v_i. Each matrix C_{v_i} is either a blocky matrix or the difference of two blocky matrices. Since the γ_2 norm of a Blocky matrix is at most 1, it follows that $\|C_{v_i}\|_{\gamma_2} \leq 2$. Since γ_2 is an algebra norm (i.e., $\|X \circ Y\|_{\gamma_2} \leq \|X\|_{\gamma_2}\|Y\|_{\gamma_2}$), we have $\|M_\ell\|_{\gamma_2} \leq 2^d$. Note that $A = \sum M_\ell$ where the sum is over all the leaves where the protocol outputs 1. Hence,

$$\|A\|_{\gamma_2} \leq 4^d. \tag{7}$$

An *unambiguous nondeterministic protocol with access to equality oracle* is a collection of 2^m deterministic equality-oracle protocols, each with depth at most d, such that on every input, at most one of them returns 1. The cost of such a protocol is $m + d$. Consider such a protocol for a Boolean matrix A, and let $A_1, \ldots, A_{2^m}$ be the Boolean matrices computed by the 2^m deterministic equality-oracle protocols. We must have $A = \sum_{i=1}^{2^m} A_i$, and in particular, by Equation (7), we have

$$\|A\|_{\gamma_2} \leq \sum_{i=1}^{2^m} \|A_i\|_{\gamma_2} \leq 2^m \times 4^d = 2^{m+2d}.$$

We denote by $\mathrm{UP}^{\textsc{eq}}(A)$, the smallest cost of an unambiguous nondeterministic equality-oracle protocol for A. We conclude

$$\frac{1}{2} \log \|A\|_{\gamma_2} \leq \mathrm{UP}^{\textsc{eq}}(A) \leq \mathrm{D}^{\textsc{eq}}(A). \tag{8}$$

2.3 Fourier Analysis of $\mathbb{Z}_p^k$

This section gives a basic overview of Fourier analysis on the finite Abelian group $G \coloneqq \mathbb{Z}_p^k$ for $p, k \in \mathbb{N}$. Consider the Hilbert space $L^2(G)$ with the inner product of two functions $f, g : G \to \mathbb{C}$ defined by

$$\langle f, g \rangle = \sum_{x \in G} f(x)\overline{g(x)}.$$

The inner product defines the norm $\|f\|_2 = \sqrt{\langle f, f \rangle}$.

Consider the principal p-th root of unity $\omega \coloneqq e^{2\pi i/p}$. For every element $a = (a_1, \ldots, a_k) \in \mathbb{Z}_p^k$, define the corresponding Fourier character $\chi_a : G \to \mathbb{C}$ as

$$\chi_a(x) = \omega^{\sum_{j=1}^k a_j x_j}.$$

The Fourier characters form an orthogonal basis for $L^2(G)$:

$$\langle \chi_a, \chi_b \rangle = \sum_{x \in G} \chi_{a-b}(x) = \begin{cases} |G| & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}.$$

Therefore, every function $f : G \to \mathbb{C}$ has a unique expansion

$$f = \sum_{a \in G} \widehat{f}(a)\chi_a,$$

where

$$\widehat{f}(a) = \frac{1}{|G|}\langle f, \chi_a \rangle.$$

It follows from the orthogonality of the Fourier characters that for every $f : G \to \mathbb{C}$,

$$\sum_{x \in G} |f(x)|^2 = |G| \sum_{a \in G} |\hat{f}(a)|^2. \tag{9}$$

This identity is called *Parseval's identity*.

3 Overview of the Proof of the Main Theorem

Let $1 \le q \le p$, and let M be the $([q] \times \mathbb{Z}_p) \times ([q] \times \mathbb{Z}_p)$ Boolean matrix defined as $M[(x, x'), (y, y')] = 1$ iff $xy = x' + y'$. Note that $M[(x, x'), (y, y')] = \mathtt{PL}_{\mathbb{Z}_p}[(x, -x'), (y, y')]$, and thus M is just a row permutation of $\mathtt{PL}_{\mathbb{Z}_p}$. Therefore, $\|M\|_{\mathrm{Tr}} = \|\mathtt{PL}_{\mathbb{Z}_p}\|_{\mathrm{Tr}}$.

Let $\sigma_1 \ge \ldots \ge \sigma_N$ be the singular values of M. Since M is a real symmetric matrix and every row of M contains exactly q ones, the largest eigenvalue of M is $\sigma_1 = q$, which corresponds to the all-1 eigenvector. If M were a "pseudo-random" matrix in the sense that all of its non-principal eigenvalues were small (i.e., $\sigma_2 < q^{1-\epsilon}$), then one could easily show that the trace norm of M is large. Indeed, the Frobenius norm of M is equal to

$$\sqrt{\sum_{(x,x'),(y,y')} M[(x, x'), (y, y')]^2} = \sqrt{qN},$$

therefore

$$\|M\|_{\mathrm{Tr}} \ge \sum_{i=2}^{N} \sigma_i \ge \frac{\sum_{i=2}^N \sigma_i^2}{\sigma_2} = \frac{qN - q^2}{\sigma_2} = \Omega\left(\frac{qN}{\sigma_2}\right). \tag{10}$$

However, we cannot expect M to be pseudo-random since pseudo-random matrices have large randomized communication complexity and this is not the case for M.

To prove a lower bound for $\|M\|_{\mathrm{Tr}}$, there is nothing special about removing only the largest singular value in Equation (10). One can take any subspace $W \subseteq \mathbb{R}^N$ and apply Equation (4) to the orthogonal projection of M to W. More precisely, let $P_W : \mathbb{R}^N \to \mathbb{R}^N$ be the orthogonal projection from $\mathbb{R}^N$ to W. By Equation (4), we have

$$\|M\|_{\mathrm{Tr}} \geq \|P_W^*\|\|M\|_{\mathrm{Tr}}\|P_W\| \geq \|P_W^* M P_W\|_{\mathrm{Tr}} \geq \frac{\|P_W^* M P_W\|_F^2}{\|P_W^* M P_W\|}.$$

Taking W as the orthogonal complement of the principal eigenvector of M yields Equation (10). The natural choice to strengthen this lower bound is to take W as the span of the eigenvectors of M that correspond to small eigenvalues. Dropping the first $k-1$ largest eigenvalues will result in the lower bound $\|M\|_{\mathrm{Tr}} \geq \frac{\sum_{i=k}^N \sigma_k^2}{\sigma_k}$. If a non-negligible mass of $\|M\|_F^2$ is on the tail $\sum_{i=k}^N \sigma_k^2$ for some $\sigma_k < q^{1-\epsilon}$, then this approach provides a strong lower bound for $\|M\|_{\mathrm{Tr}}$.

Unfortunately, the direct application of this method requires determining the eigenvectors and eigenvalues of M, which seems difficult. To circumvent this difficult task, we employ tools from Fourier analysis and show that there is a linear span of some Fourier characters $W \subseteq \mathbb{R}^N$ such that $\|P_W^* M P_W\|_F = \Omega(\|M\|_F)$ and $\|P_W^* M P_W\|$ is small.

4 Randomized Communication Complexities of PL and $\mathtt{PL}_{\mathbb{Z}_p}$

We divide the proof of Theorem 4 into two sections. In this section, we prove the upper bounds of Theorem 4 on $\mathrm{R}_0^1(\mathtt{PL}_{\mathbb{Z}_p})$ and $\mathrm{R}_0^1(\mathtt{PL})$.

▶ **Proposition 12.** *For $q \leq \sqrt{p}$, we have $\mathrm{R}_0^1(\mathtt{PL}_{\mathbb{Z}_p}) = O(\log\log p)$. For every $1 \leq q \leq p$, we have $\mathrm{R}_0^1(\mathtt{PL}) = O(\log\log p)$.*

Proof. We describe a randomized protocol that solves $\mathtt{PL}_{\mathbb{Z}_p}$ with cost $O(\log\log p)$ that never makes mistakes on inputs where $\mathtt{PL}_{\mathbb{Z}_p}$ takes value 1. The same protocol also solves PL.

Suppose Alice and Bob have inputs $(x, x'), (y, y') \in [q] \times \mathbb{Z}_p$ respectively. Since $q \leq \sqrt{p}$, we have

$$[xy + x' \equiv_p y'] \iff [xy + x' = y'] \vee [xy + x' = y' + p].$$

In the rest of the proof, we show that each of the two equations on the right-hand side can be verified with a protocol of cost at most $O(\log\log p)$ and error at most $1/6$, which then implies a protocol of cost $O(\log\log p)$ and error at most $1/3$ for the matrix $\mathtt{PL}_{\mathbb{Z}_p}$. Suppose Alice and Bob want to verify whether $xy + x' = y'$; the case for $xy + x' = y' + p$ is similar. Alice picks a uniformly random prime $\mathbf{r}$ from the set of the first $\lceil 6\log(2p) \rceil$ primes $\mathcal{P}$ and sends it to Bob. Alice and Bob exchange the values $(x \mod \mathbf{r}), (x' \mod \mathbf{r}), (y \mod \mathbf{r}), (y' \mod \mathbf{r})$ and check whether

$$(x \mod \mathbf{r})(y \mod \mathbf{r}) + (x' \mod \mathbf{r}) \equiv_{\mathbf{r}} (y' \mod \mathbf{r}),$$

or equivalently

$$xy + x' \equiv_{\mathbf{r}} y'.$$

The cost of this communication is at most $O(\log \mathbf{r}) = O(\log\log p)$. Next, we show that the probability of error (over the choice of $\mathbf{r}$) is at most $1/6$. Observe that an error can only happen when $xy + x' \neq y'$ but $xy + x' \equiv_{\mathbf{r}} y'$. We want to show that

$$\Pr_{\mathbf{r} \in \mathcal{P}}[[xy + x' \neq y'] \wedge [xy + x' \equiv_{\mathbf{r}} y']] \leq \frac{1}{6}.$$

Let $B \subseteq \mathcal{P}$ be the set of bad choices for $\mathbf{r}$, namely

$$B = \{r \in \mathcal{P} : [xy + x' \neq y'] \wedge [xy + x' \equiv_{\mathbf{r}} y']\}.$$

Suppose towards a contradiction that $|B| > \frac{|\mathcal{P}|}{6}$. Define

$$m \coloneqq \prod_{r \in B} r \geq 2^{|B|} > 2p.$$

Note that for all $r \in B$, we have $xy + x' \equiv_r y'$. By the Chinese remainder theorem, we have $xy + x' \equiv_m y'$. This implies the contradiction that $xy + x' = y'$ because $0 \leq xy + x', y' < 2p < m$. ◀

▶ Remark 13. Note that the protocol used in the proof Proposition 12 is in fact a private-coin protocol, so the bounds in Proposition 12 hold in both private-coin and public-coin models.

▶ Remark 14. Combining Proposition 12 with Equation (6), we obtain

$$\widetilde{\gamma}_2(\mathtt{PL}_{\mathbb{Z}_p}) \leq \log^{O(1)}(N). \tag{11}$$

5 Trace Norms of PL and $\mathtt{PL}_{\mathbb{Z}_p}$

This section is dedicated to proving the lower bounds on $\|\mathtt{PL}_{\mathbb{Z}_p}\|_{\mathrm{Tr}}$ and $\|\mathtt{PL}\|_{\mathrm{Tr}}$ of Theorem 4. The lower bounds on $\|\mathtt{PL}_{\mathbb{Z}_p}\|_{\gamma_2}$ and $\|\mathtt{PL}\|_{\gamma_2}$ immediately follow from Equation (5).

▶ **Lemma 15.** *For $1 \leq q \leq p$, we have*

$$\|\mathtt{PL} - \mathtt{PL}_{\mathbb{Z}_p}\|_{\mathrm{Tr}} \leq q^4.$$

In particular, if $q \leq p^{1/3}$, then

$$\|\mathtt{PL} - \mathtt{PL}_{\mathbb{Z}_p}\|_{\mathrm{Tr}} = O(pq).$$

Proof. For $(x, x'), (y, y') \in [q] \times \{0, \ldots, p-1\}$, we have

$$\mathtt{PL}_{\mathbb{Z}_p}[(x, x'), (y, y')] = 1 \text{ iff } [xy + x' = y'] \vee [xy + x' = y' + p].$$

Therefore, we can write $\mathtt{PL}_{\mathbb{Z}_p} = \mathtt{PL} + A$, where A is defined as

$$A[(x, x'), (y, y')] = 1 \text{ iff } xy + x' = y' + p.$$

Because $xy \leq q^2$ and $x' < p$, $xy + x' = y' + p$ implies $y' < q^2$. Therefore, A has at most q^4 non-zero entries. Consequently $\|A\|_{\mathrm{Tr}} \leq q^4$. ◀

By Lemma 15, to complete the proof of Theorem 4, it suffices to prove $\|\mathtt{PL}_{\mathbb{Z}_p}\|_{\mathrm{Tr}} = \Omega(pq^{9/8})$. Since we want to apply Fourier analysis to study the trace norm of $\mathtt{PL}_{\mathbb{Z}_p}$, it is more convenient to extend the rows and columns of $\mathtt{PL}_{\mathbb{Z}_p}$ to $G \coloneqq \mathbb{Z}_p^2$ by adding all-zero rows and columns. That is, we consider $M : G \times G \to \{0, 1\}$, defined as

$$M[(x, x'), (y, y')] = \begin{cases} 1 & \text{if } x, y \in [q] \text{ and } xy \equiv_p x' + y' \\ 0 & \text{otherwise} \end{cases}.$$

This definition of M is slightly different from the one used in the proof overview, but all of the properties we want still hold. For $x, y \in [q]$, we have $M[(x, x'), (y, y')] = \mathtt{PL}_{\mathbb{Z}_p}[(x, -x'), (y, y')]$, and M is zero on the other entries. In other words, M is obtained

from $\mathtt{PL}_{\mathbb{Z}_p}$ by first permuting the rows according to the change of variable $x' \to -x'$, then adding several all-zero rows and columns. These operations do not change the matrix's trace, Frobenius, and spectral norm, and in particular,

$$\|\mathtt{PL}_{\mathbb{Z}_p}\|_{\mathrm{Tr}} = \|M\|_{\mathrm{Tr}}.$$

For $(\alpha, \beta) \in G$, let $\chi_{\alpha,\beta} : G \to \mathbb{C}$ denote the corresponding character in $\widehat{G}$, defined as $\chi_{\alpha,\beta} : (x, x') \mapsto \omega^{\alpha x + \beta x'}$ where $\omega = e^{2\pi i/p}$.

Let $S \subseteq \mathbb{Z}_p$, and π_S be the $G \times G$ matrix corresponding to the orthogonal projection from $L^2(G)$ to the span of $\chi_{\alpha,\beta}$ for $(\alpha, \beta) \in \mathbb{Z}_p \times S$. That is, for $f : G \to \mathbb{C}$,

$$\pi_S f = \sum_{\alpha \in \mathbb{Z}_p} \sum_{\beta \in S} \widehat{f}(\alpha, \beta) \chi_{\alpha,\beta}.$$

Denote $M_S \coloneqq \pi_S^* M \pi_S$. Since π_S is an orthogonal projection, we have $\pi_S = \pi_S^*$ and $\|\pi_S\| = \|\pi_S^*\| \leq 1$, and therefore,

$$\|M_S\|_{\mathrm{Tr}} = \|\pi_S^* M \pi_S\|_{\mathrm{Tr}} \leq \|\pi_S^*\| \|M\|_{\mathrm{Tr}} \|\pi_S\| \leq \|M\|_{\mathrm{Tr}}.$$

Hence, we can use Equation (4) to obtain a lower bound for $\|M\|_{\mathrm{Tr}}$:

$$\|M\|_{\mathrm{Tr}} \geq \|M_S\|_{\mathrm{Tr}} \geq \frac{\|M_S\|_F^2}{\|M_S\|}.$$

First, we determine the value of $\|M_S\|_F$.

▶ **Lemma 16.** *For any $S \subseteq \mathbb{Z}_p$, $\|M_S\|_F = q\sqrt{|S \cap (-S)|}$.*

Proof. Since $\frac{1}{\sqrt{|G|}}\chi_{\alpha,\beta}$'s form an orthonormal basis for $L^2(G)$, for every matrix $B \in \mathbb{C}^{G \times G}$, we have

$$\|B\|_F^2 = \frac{1}{|G|^2} \sum_{(\alpha,\beta),(\alpha',\beta') \in G} |\langle B\chi_{\alpha,\beta}, \chi_{\alpha',\beta'}\rangle|^2. \tag{12}$$

For every $\alpha, \alpha', \beta, \beta' \in \mathbb{Z}_p$, we have

$$\langle M_S \chi_{\alpha,\beta}, \chi_{\alpha',\beta'}\rangle = \langle M\pi_S\chi_{\alpha,\beta}, \pi_S\chi_{\alpha',\beta'}\rangle = \begin{cases} \langle M\chi_{\alpha,\beta}, \chi_{\alpha',\beta'}\rangle & \text{if } \beta, \beta' \in S \\ 0 & \text{otherwise} \end{cases}$$

and therefore, by Equation (12),

$$\|M_S\|_F^2 = \frac{1}{|G|^2} \sum_{(\alpha,\beta),(\alpha',\beta') \in \mathbb{Z}_p \times S} |\langle M\chi_{\alpha,\beta}, \chi_{\alpha',\beta'}\rangle|^2. \tag{13}$$

For $\beta \in S$, define the matrix $F_\beta \in \mathbb{C}^{p \times p}$ as

$$F_\beta(\alpha, \alpha') = \sum_{x,y \in [q]} \omega^{\alpha x + \alpha' y + \beta xy}. \tag{14}$$

Let $\alpha, \alpha' \in \mathbb{Z}_p$ and $\beta, \beta' \in S$. We have

$$\begin{aligned}
\langle M\chi_{\alpha,\beta}, \chi_{\alpha',\beta'}\rangle &= \sum_{x,y\in\mathbb{Z}_p}\sum_{x',y'\in\mathbb{Z}_p} M[(y,y'),(x,x')]\chi_{\alpha,\beta}(x,x')\overline{\chi_{\alpha',\beta'}(y,y')} \\
&= \sum_{x,y\in[q]}\sum_{x',y'\in\mathbb{Z}_p} M[(y,y'),(x,x')]\chi_{\alpha,\beta}(x,x')\overline{\chi_{\alpha',\beta'}(y,y')} \\
&= \sum_{x,y\in[q]}\sum_{y'\in\mathbb{Z}_p} \chi_{\alpha,\beta}(x,xy-y')\overline{\chi_{\alpha',\beta'}(y,y')} \\
&= \sum_{x,y\in[q]}\sum_{y'\in\mathbb{Z}_p} \omega^{\alpha x+\beta(xy-y')-\alpha' y-\beta' y'} \\
&= \sum_{x,y\in[q]} \omega^{\alpha x-\alpha' y+\beta xy}\sum_{y'\in\mathbb{Z}_p} \omega^{-(\beta+\beta')y'} \\
&= \begin{cases} pF_\beta(\alpha,-\alpha') & \text{if } \beta=-\beta' \\ 0 & \text{otherwise} \end{cases}.
\end{aligned}$$

Combining this with Equation (13) gives

$$\|M_S\|_F^2 = \frac{p^2}{|G|^2}\sum_{\alpha,\alpha'\in\mathbb{Z}_p}\sum_{\beta\in S\cap(-S)}|F_\beta(\alpha,-\alpha')|^2 = \frac{1}{|G|}\sum_{\beta\in S\cap(-S)}\|F_\beta\|_F^2. \tag{15}$$

Furthermore,

$$\begin{aligned}
\|F_\beta\|_F^2 &= \sum_{\alpha,\alpha'\in\mathbb{Z}_p}\sum_{x,y\in[q]}\omega^{\alpha x+\alpha' y+\beta xy}\sum_{x',y'\in[q]}\omega^{-(\alpha x'+\alpha' y'+\beta x'y')} \\
&= \sum_{\alpha,\alpha'\in\mathbb{Z}_p}\sum_{x,y,x',y'\in[q]}\omega^{\alpha(x-x')+\alpha'(y-y')+\beta(xy-x'y')} \\
&= \sum_{x,y,x',y'\in[q]}\omega^{\beta(xy-x'y')}\sum_{\alpha,\alpha'\in\mathbb{Z}_p}\omega^{\alpha(x-x')+\alpha'(y-y')}.
\end{aligned}$$

The inner sum is zero unless $x=x'$ and $y=y'$, in which case the inner sum is evaluated to p^2. Thus, for every β, we have $\|F_\beta\|_F^2 = q^2p^2$. We conclude that

$$\|M_S\|_F^2 = \frac{|S\cap(-S)|q^2p^2}{|G|} = q^2|S\cap(-S)|. \qquad \blacktriangleleft$$

Next, we turn to the upper bound of the spectral norm of M_S.

▶ **Lemma 17.** *There is a set $S\subseteq\mathbb{Z}_p$, closed under negation and of size $|S|\geq p/2$, such that $\|M_S\|\leq 2q^{7/8}$.*

Proof. We have

$$\|M_S\| = \max_{\substack{f,g:G\to\mathbb{C}\\ \|f\|_2=\|g\|_2=1}}\langle M_S f,g\rangle = \max_{\substack{f,g:G\to\mathbb{C}\\ \|f\|_2=\|g\|_2=1}}\langle M\pi_S f,\pi_S g\rangle.$$

Define $\widehat{f}_\beta,\widehat{g}_\beta\in\mathbb{C}^p$ as $\widehat{f}_\beta(\alpha):=\widehat{f}(\alpha,\beta)$ and $\widehat{g}_\beta(\alpha):=\widehat{g}(-\alpha,-\beta)$ for each $\alpha\in\mathbb{Z}_p$. Recalling the definition of F_β in Equation (14), we have

$$\begin{aligned}
\langle M\pi_S f, \pi_S g\rangle &= \sum_{\beta,\beta'\in S}\sum_{\alpha,\alpha'\in\mathbb{Z}_p} \widehat{f}(\alpha,\beta)\overline{\widehat{g}(\alpha',\beta')}\langle M\chi_{\alpha,\beta}, \chi_{\alpha',\beta'}\rangle \\
&= \sum_{\beta\in S\cap(-S)}\sum_{\alpha,\alpha'\in\mathbb{Z}_p} \widehat{f}(\alpha,\beta)\overline{\widehat{g}(\alpha',-\beta)}\langle M\chi_{\alpha,\beta}, \chi_{\alpha',-\beta}\rangle \\
&= p\sum_{\beta\in S}\sum_{\alpha,\alpha'\in\mathbb{Z}_p} \widehat{f}(\alpha,\beta)\overline{\widehat{g}(\alpha',-\beta)}F_\beta(\alpha,-\alpha') \\
&= p\sum_{\beta\in S}\sum_{\alpha,\alpha'\in\mathbb{Z}_p} \widehat{f}(\alpha,\beta)\overline{\widehat{g}(-\alpha',-\beta)}F_\beta(\alpha,\alpha') \\
&= p\sum_{\beta\in S}\sum_{\alpha,\alpha'\in\mathbb{Z}_p} \widehat{f}(\alpha,\beta)\overline{\widehat{g}(-\alpha',-\beta)}F_\beta(\alpha',\alpha) \\
&= p\sum_{\beta\in S}\sum_{\alpha,\alpha'\in\mathbb{Z}_p} \widehat{f}_\beta(\alpha)\widehat{g}_\beta(\alpha')F_\beta(\alpha',\alpha) \\
&= p\sum_{\beta\in S}\langle F_\beta\widehat{f}_\beta, \widehat{g}_\beta\rangle,
\end{aligned}$$

where at the third equality, we used the negation-closed property of S. By the definition of spectral norm and Cauchy-Schwarz inequality,

$$\begin{aligned}
|\langle M\pi_S f, \pi_S g\rangle| &\le p\sum_{\beta\in S}|\langle F_\beta\widehat{f}_\beta, \widehat{g}_\beta\rangle| \le p\sum_{\beta\in S}\|F_\beta\|\|\widehat{f}_\beta\|_2\|\widehat{g}_\beta\|_2 \\
&\le p\max_{\beta\in S}\|F_\beta\|\sqrt{\sum_{\beta\in S}\|\widehat{f}_\beta\|_2^2}\sqrt{\sum_{\beta\in S}\|\widehat{g}_\beta\|_2^2} \le \frac{p}{|G|}\max_{\beta\in S}\|F_\beta\|,
\end{aligned}$$

where the last inequality follows from Parseval's identity Equation (9) and $\|f\|_2 = \|g\|_2 = 1$:

$$\sum_{\beta\in S}\|\widehat{f}_\beta\|_2^2 \le \sum_{\beta\in\mathbb{Z}_p}\|\widehat{f}_\beta\|_2^2 = \sum_{(\alpha,\beta)\in G}|\widehat{f}(\alpha,\beta)|^2 = \frac{1}{|G|}\sum_{(x,y)\in G}|f(x,y)|^2 = \frac{1}{|G|}.$$

Next, we upper-bound the spectral norm of F_β using the 4th moment of singular values:

$$\begin{aligned}
\|F_\beta\|^4 &\le \|F_\beta\|_{S_4}^4 = \mathrm{Tr}\big(F_\beta F_\beta^* F_\beta F_\beta^*\big) \\
&= \sum_{\alpha_1,\alpha_1',\alpha_2,\alpha_2'\in\mathbb{Z}_p} F_\beta(\alpha_1,\alpha_1')\overline{F_\beta(\alpha_1,\alpha_2')}F_\beta(\alpha_2,\alpha_2')\overline{F_\beta(\alpha_2,\alpha_1')} \\
&= \sum_{\substack{\alpha_1,\alpha_1'\\ \alpha_2,\alpha_2'}}\sum_{\substack{x_1,\dots,x_4\\ y_1,\dots,y_4}} \omega^{\alpha_1(x_1-x_2)+\alpha_2(x_3-x_4)+\alpha_1'(y_1-y_4)+\alpha_2'(y_3-y_2)}\omega^{\beta(x_1y_1-x_2y_2+x_3y_3-x_4y_4)} \\
&= \sum_{\substack{x_1,\dots,x_4\\ y_1,\dots,y_4}}\sum_{\substack{\alpha_1,\alpha_1'\\ \alpha_2,\alpha_2'}} \omega^{\alpha_1(x_1-x_2)+\alpha_2(x_3-x_4)+\alpha_1'(y_1-y_4)+\alpha_2'(y_3-y_2)}\omega^{\beta(x_1y_1-x_2y_2+x_3y_3-x_4y_4)}.
\end{aligned}$$

The inner sum is zero unless $x_1 = x_2$, $x_3 = x_4$, $y_1 = y_4$ and $y_2 = y_3$. This simplifies $\|F_\beta\|_{S_4}^4$ to

$$\|F_\beta\|_{S_4}^4 = p^4\sum_{x,y,x',y'\in[q]}\omega^{\beta(xy-xy'+x'y'-x'y)} = p^4 r(\beta),$$

where

$$r(\beta) \coloneqq \sum_{\vec{u}\in[q]^4}\omega^{\beta\phi(\vec{u})} \qquad \text{with} \qquad \phi(u_1,u_2,u_3,u_4) \coloneqq u_1u_2 - u_1u_4 + u_3u_4 - u_3u_2.$$

For every $z \in \mathbb{Z}_p$ and $y \in \mathbb{Z}_p \setminus \{0\}$, we have $\Pr_{x\in[q]}[xy \equiv_p z] \in \{0, 1/q\}$. Note that the event $\{\phi(\vec{u}) \equiv_p \phi(\vec{v})\}$ is equivalent to $\{u_1(u_2 - u_4) \equiv_p z\}$, where $z = v_1v_2 - v_1v_4 + v_3v_4 - v_3v_2 - u_3u_4 + u_3u_2$. Consider uniform independent random variables $\vec{u}, \vec{v} \in [q]^4$. Conditioned on $u_2 \neq u_4$, which happens with probability $1 - 1/q$, the probability that $u_1(u_2 - u_4) \equiv_p z$ is at most $1/q$. Therefore,

$$\Pr[\phi(\vec{u}) \equiv_p \phi(\vec{v})] \leq \left(1 - \frac{1}{q}\right) \times \frac{1}{q} + \frac{1}{q} \times 1 \leq \frac{2}{q},$$

implying that $|\{(\vec{u}, \vec{v}) : \phi(\vec{u}) \equiv_p \phi(\vec{v})\}| \leq 2q^7$. Hence

$$\mathop{\mathbb{E}}_{\beta} |r(\beta)|^2 = \mathop{\mathbb{E}}_{\beta} \left[\sum_{\vec{u},\vec{v}\in[q^4]} \omega^{\beta(\phi(\vec{u})-\phi(\vec{v}))} \right] = \sum_{\vec{u},\vec{v}} \mathop{\mathbb{E}}_{\beta} \left[\omega^{\beta(\phi(\vec{u})-\phi(\vec{v}))}\right] = \sum_{\vec{u},\vec{v}} \mathbf{1}_{\{\phi(\vec{u})\equiv_p\phi(\vec{v})\}} \leq 2q^7.$$

From the above inequality, for $t \coloneqq 2q^{7/2}$, by Markov's inequality we have

$$\Pr_{\beta}[|r(\beta)| \geq t] \leq \frac{2q^7}{t^2} = \frac{1}{2}.$$

As $\beta\phi(u_1, u_2, u_3, u_4) = -\beta\phi(u_1, u_4, u_3, u_2)$, we have $r(\beta) = r(-\beta)$ for any β, and so

$$S \coloneqq \{\beta \in \mathbb{Z}_p : |r(\beta)| < t\}$$

is a subset of $\mathbb{Z}_p$ closed under negation with $|S| \geq p/2$. Therefore,

$$\|M_S\| \leq \frac{p}{|G|} \max_{\beta\in S} \|F_\beta\| \leq \frac{p}{|G|} \max_{\beta\in S} \|F_\beta\|_{S_4} \leq \frac{p}{|G|} \max_{\beta\in S} \{p|r(\beta)|^{1/4}\} < t^{1/4} \leq 2q^{7/8}. \qquad \blacktriangleleft$$

By combining Lemma 16 and Lemma 17, we conclude that

$$\|M\|_{\mathrm{Tr}} \geq \|M_S\|_{\mathrm{Tr}} \geq \frac{\|M_S\|_F^2}{\|M_S\|} \geq \frac{q^2 \times p/2}{2q^{7/8}} = \Omega(pq^{9/8}).$$

6 Concluding Remarks

We showed the existence of Boolean matrices $M_{N\times N}$ with $\|M\|_{\gamma_2} \geq \Omega(N^{1/32})$ and $\mathrm{R}(M) \leq \mathrm{R}^1_0(M) \leq O(\log\log N)$, displaying a double exponential separation between γ_2 norm and randomized communication complexity. We did not attempt to optimize the power of N in the lower bound, and there is no reason to suspect that $1/32$ is the best possible.

▶ **Question 3.** *What is the largest c such that there exist Boolean matrices $M_{N\times N}$ with $\mathrm{R}(M) \leq O(\log\log N)$ and $\|M\|_{\gamma_2} \geq \Omega(N^c)$?*

It is also natural to ask the analogue of Question 3 regarding the approximate γ_2 norm.

▶ **Question 4.** *What is the largest c such that there exist Boolean matrices $M_{N\times N}$ with $\widetilde{\gamma}_2(M) \leq \mathrm{polylog}(N)$ and $\|M\|_{\gamma_2} \geq \Omega(N^c)$?*

We remark that in Question 4, one cannot hope to obtain a lower bound stronger than $\Omega(N^{1/2})$ as $\|M\|_{\gamma_2} \leq \widetilde{\gamma}_2(M) + O(\sqrt{N})$ for all M (see [12, Lemma 15]).

Whether or not the upper bounds in Question 3 and Question 4 can be improved is also an interesting open problem. As we discussed in Section 1.3, there are Boolean matrices $M_{N\times N}$ with $\mathrm{R}(M) = O(1)$ but $\|M\|_{\gamma_2} = \mathrm{polylog}(N)$.

▶ **Question 5.** *Is there a Boolean matrix* $M_{N\times N}$ *with* $\mathrm{R}(M) = O(1)$ *and* $\|M\|_{\gamma_2} = N^{\Omega(1)}$*?*

We could not overrule the possibility that PL and IIP are such examples. Nevertheless, we make the following conjecture.

▶ **Conjecture 18.** $\mathrm{R}(\mathtt{PL}) = \omega(1)$.

Another intriguing question is about the relationship between γ_2 norm and the *unbounded error* randomized communication complexity, denoted by $\mathrm{U}(\cdot)$. It is well-known [15] that $\mathrm{U}(M) = \log \mathrm{rank}_{\pm}(M) \pm O(1)$ where $\mathrm{rank}_{\pm}(\cdot)$ denotes the sign-rank a.k.a. dimension complexity. The reader is referred to [6] for the definitions of $\mathrm{U}(\cdot)$ and $\mathrm{rank}_{\pm}$. It is natural to ask whether one can obtain an upper bound on sign-rank based solely on γ_2 norm. In other words, the following conjecture is intriguing.

▶ **Conjecture 19.** *Suppose* $\|M\|_{\gamma_2} = O(1)$*. Then* $\mathrm{rank}_{\pm}(M) = O(1)$.

A viable approach to settle the above question in the positive is by using the parameter $\mathrm{D}^{\mathrm{EQ}}(\cdot)$. Hatami et al. [6] showed that if $\mathrm{D}^{\mathrm{EQ}}(M) = O(1)$, then $\mathrm{rank}_{\pm}(M) = O(1)$. On the other hand, the following was conjectured in [5], which, if true, would imply Conjecture 19.

▶ **Conjecture 20 ([5]).** *Suppose* $\|M\|_{\gamma_2} = O(1)$*. Then* $\mathrm{D}^{EQ}(M) = O(1)$.

In the special case where M is an XOR function, it is shown in [6] that Conjecture 20 is true. The authors show that this follows from Green-Sanders' quantitative version of Cohen's idempotent theorem [4].

References

1. Shalev Ben-David, Adam Bouland, Ankit Garg, and Robin Kothari. Classical lower bounds from quantum upper bounds. *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 339–349, 2018.
2. Arkadev Chattopadhyay, Shachar Lovett, and Marc Vinyals. Equality alone does not simulate randomness. In *34th Computational Complexity Conference (CCC 2019)*, 2019.
3. Alexander Edmonds, Aleksandar Nikolov, and Jonathan Ullman. The power of factorization mechanisms in local and central differential privacy. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 425–438, New York, NY, USA, June 2020. Association for Computing Machinery.
4. Ben Green and Tom Sanders. Boolean functions with small spectral norm. *Geometric and Functional Analysis*, 18(1):144–162, 2008.
5. Lianna Hambardzumyan, Hamed Hatami, and Pooya Hatami. Dimension-free bounds and structural results in communication complexity. *Israel J. Math.*, 2022. doi:10.1007/s11856-022-2365-8.
6. Hamed Hatami, Pooya Hatami, William Pires, Ran Tao, and Rosie Zhao. Lower bound methods for sign-rank and their limitations. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms a*, volume 245, pages 22:1–22:24, 2022.
7. Monika Henzinger and Jalaj Upadhyay. Constant matters: Fine-grained complexity of differentially private continual observation using completely bounded norms. *arXiv preprint*, 2022. arXiv:2202.11205.
8. Wei Huang, Yaoyun Shi, Shengyu Zhang, and Yufan Zhu. The communication complexity of the Hamming distance problem. *Inform. Process. Lett.*, 99(4):149–153, 2006.
9. Hartmut Klauck. Lower bounds for quantum communication complexity. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 288–297. IEEE, 2001.
10. Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, Cambridge, 1997.

11 Troy Lee, Adi Shraibman, and Robert Špalek. A direct product theorem for discrepancy. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 71–80, 2008. doi:10.1109/CCC.2008.25.

12 Nati Linial and Adi Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures & Algorithms*, 34(3):368–394, 2009.

13 Jiri Matousek, Aleksandar Nikolov, and Kunal Talwar. Factorization norms and hereditary discrepancy. *arXiv preprint*, 2014. arXiv:1408.1376.

14 Shanmugavelayutham Muthukrishnan and Aleksandar Nikolov. Optimal private halfspace counting via discrepancy. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 1285–1292, 2012.

15 Ramamohan Paturi and Janos Simon. Probabilistic communication complexity. *Journal of Computer and System Sciences*, 33(1):106–123, 1986.

16 Toniann Pitassi, Morgan Shirley, and Adi Shraibman. The strength of equality oracles in communication. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 89:1–89:19, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITCS.2023.89.

17 Alexander A Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, 2003.

18 Suhail Sherif. *Communication Complexity and Quantum Optimization Lower Bounds via Query Complexity*. PhD thesis, Tata Institute of Fundamental Research, Mumbai, 2021.

19 József Solymosi. Incidences and the spectra of graphs. In *Combinatorial number theory and additive group theory*, pages 299–314. Springer, 2009.