

Polytopes with Bounded Integral Slack Matrices Have Sub-Exponential Extension Complexity

Sally Dong*

Thomas Rothvoss†

March 19, 2024

Abstract

We show that any bounded integral function $f : A \times B \mapsto \{0, 1, \dots, \Delta\}$ with rank r has deterministic communication complexity $\Delta^{O(\Delta)} \cdot \sqrt{r} \cdot \log r$, where the rank of f is defined to be the rank of the $A \times B$ matrix whose entries are the function values. As a corollary, we show that any n -dimensional polytope that admits a slack matrix with entries from $\{0, 1, \dots, \Delta\}$ has extension complexity at most $\exp(\Delta^{O(\Delta)} \cdot \sqrt{n} \cdot \log n)$.

1 Introduction

In classical communication complexity, two players, Alice and Bob, are given a Boolean function $f : A \times B \mapsto \{0, 1\}$, as well as separate inputs $a \in A$ and $b \in B$, and wish to compute $f(a, b)$ while minimizing the total amount of communication. Alice and Bob have unlimited resources for pre-computations and agree on a deterministic *communication protocol* to compute f before receiving their respective inputs. The *length* of their protocol is defined to be the maximum number of bits exchanged over all possible inputs. The *deterministic communication complexity* of f , denoted by $CC^{\det}(f)$, is the minimum length of a protocol to compute f .

A major open problem in communication complexity is the *log-rank conjecture* proposed by Lovász and Saks [LS88], which asks if $CC^{\det}(f) \leq (\log r)^{O(1)}$ for all Boolean functions f of rank r , where the *rank* of a two-party function f on $A \times B$ is defined to be the rank of the matrix $\mathbf{M} \in \mathbb{R}^{A \times B}$ with $\mathbf{M}_{a,b} = f(a, b)$ for all $(a, b) \in A \times B$. The best upper bound currently known is due to Lovett [Lov16], who showed $CC^{\det}(f) \leq O(\sqrt{r} \log r)$ using discrepancy theory techniques.

In this work, we obtain similar deterministic communication complexity bounds for a larger class of functions:

Theorem 1 (Main result, communication complexity). *Let $f : A \times B \mapsto \{0, 1, \dots, \Delta\}$ be a bounded integral function of rank r . Then there exists a deterministic communication protocol to compute f with length at most $\Delta^{O(\Delta)} \cdot \sqrt{r} \cdot \log r$ bits.*

The function f can be directly viewed as a non-negative matrix $\mathbf{M} \in \{0, \dots, \Delta\}^{A \times B}$ of rank r . We use the matrix representation exclusively in the remainder of this paper. Let us adopt the convention that a *rectangle* in \mathbf{M} is a (non-contiguous) submatrix $\mathbf{M}[A', B']$ indexed by some $A' \subseteq A$ and $B' \subseteq B$. A rectangle is *monochromatic* with color i if all entries in the rectangle have value i .

*University of Washington. sallyqd@uw.edu.

†University of Washington. rothvoss@uw.edu. Supported by NSF CAREER grant 1651861, NSF grant 2318620 and a David & Lucile Packard Foundation Fellowship.

The *non-negative rank* of a non-negative matrix \mathbf{M} , denoted by $\text{rank}_+(\mathbf{M})$, is defined as the minimum r such that \mathbf{M} can be written as the sum of r non-negative rank-1 matrices, or equivalently, as $\mathbf{M} = \mathbf{UV}$ for non-negative matrices $\mathbf{U} \in \mathbb{R}_{\geq 0}^{A \times r}$ and $\mathbf{V} \in \mathbb{R}_{\geq 0}^{r \times B}$. It is straightforward to see $\text{rank}_+(\mathbf{M}) \leq 2^{CC^{\det}(\mathbf{M})}$ (c.f. Rao and Yehudayoff [RY20], Chapter 1, Theorem 1.6): The *protocol tree* to compute \mathbf{M} has at most $2^{CC^{\det}(\mathbf{M})}$ leaves, each corresponding to a monochromatic rectangle of \mathbf{M} . These rectangles are disjoint over all leaves, and their union is \mathbf{M} . Since a monochromatic rectangle is a non-negative matrix of rank 0 or 1, we conclude that \mathbf{M} can be written as a sum of at most $2^{CC^{\det}(\mathbf{M})}$ rank-1 non-negative matrices. The *positive semidefinite rank* of \mathbf{M} , denoted by $\text{rank}_{\text{psd}}(\mathbf{M})$, generalizes non-negative rank and has an analogous relationship to quantum communication complexity [FGP⁺15]. It is defined as the minimum r such that there are positive semidefinite matrices $\mathbf{U}_1, \dots, \mathbf{U}_A$ and $\mathbf{V}_1, \dots, \mathbf{V}_B$ of dimension $r \times r$ satisfying $\mathbf{M}_{i,j} = \text{Tr}(\mathbf{U}_i \mathbf{V}_j)$ for all i, j ; trivially, $\text{rank}_{\text{psd}}(\mathbf{M}) \leq \text{rank}_+(\mathbf{M})$. Barvinok [Bar12] showed that if \mathbf{M} has at most k distinct entries, as is the setting studied in this paper, then $\text{rank}_{\text{psd}}(\mathbf{M}) \leq \binom{k-1+\text{rank}(\mathbf{M})}{k-1}$.

Non-negative rank brings us to a beautiful connection with *extension complexity*, introduced in the seminal work of Yannakakis [Yan88] in the context of writing combinatorial optimization problems as linear programs. The *extension complexity* of a polytope \mathcal{P} , denoted $\text{xc}(\mathcal{P})$, is defined as the minimum number of facets of some higher dimensional polytope \mathcal{Q} (its *extended formulation*) such that there exists a linear projection of \mathcal{Q} to \mathcal{P} .

A foundational theorem from [Yan88] states that $\text{xc}(\mathcal{P}) = \text{rank}_+(\mathbf{S})$, where \mathbf{S} is the *slack matrix* of \mathcal{P} , defined as follows: Suppose \mathcal{P} has facets \mathcal{F} and vertices \mathcal{V} . Then \mathbf{S} is a non-negative $\mathcal{F} \times \mathcal{V}$ matrix, where the (f, v) -entry indexed by facet $f \in \mathcal{F}$ defined by the halfspace $a^\top x \leq b$ and vertex $v \in \mathcal{V}$ has value $\mathbf{S}_{f,v} = b - a^\top v$. (A facet may be defined by many equivalent halfspaces, and therefore the slack matrix is not unique; the result holds for all valid slack matrices.) In fact, [Yan88] showed that a factorization of \mathbf{S} with respect to non-negative rank gives an extended formulation of \mathcal{P} and vice versa. For a comprehensive preliminary survey, see [CCZ10]. A number of breakthrough results in extended complexity in recent years emerged from lower-bounding the non-negative rank of the slack matrix for specific polytopes, such as the TSP, CUT, and STABLE-SET polytopes by Fiorini-Massar-Pokutta-Tiwary-De Wolf [FMP⁺15], and the PERFECT-MATCHING polytope by Rothvoss [Rot17].

Connecting extension complexity to deterministic communication complexity via non-negative rank, we have:

Corollary 2 (Main result, extension complexity). *Let \mathcal{P} be a n -dimensional polytope that admits slack matrix \mathbf{S} . Suppose the entries of \mathbf{S} are integral and bounded by Δ . Then the extension complexity of \mathcal{P} is at most $\exp(\Delta^{O(\Delta)} \cdot \sqrt{n} \cdot \log n)$.*

Proof. Since \mathcal{P} is n -dimensional, we know \mathbf{S} has rank at most n , as there are at most n linearly independent vertices of \mathcal{P} , and the slack of a vertex with respect to all the facets is a linear function. We combine the theorem of [Yan88] and Theorem 1 for the overall conclusion. \square

Our proof follows the approach discussed in the note of Rothvoss [Rot14] simplifying Lovett's result. The following lemma shows that there are indeed concrete polytopes for which our result applies:

Lemma 3. *Suppose there are polytopes $\mathcal{P} \subseteq \mathcal{Q} \subseteq \mathbb{R}^n$ where $\mathcal{P} = \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_v\}$ and $\mathcal{Q} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ with $\mathbf{A} \in \mathbb{R}^{f \times n}$. Suppose the partial slack matrix $\mathbf{S} \in \mathbb{R}^{f \times v}$ with $\mathbf{S}_{i,j} = \mathbf{b}_i - \mathbf{A}_i \mathbf{x}_j$ for $i \in [f], j \in [v]$ is integral and bounded by Δ . Then there exists a polytope \mathcal{K} with extension complexity at most $\exp(\Delta^{O(\Delta)} \cdot \sqrt{n} \cdot \log n)$ so that $\mathcal{P} \subseteq \mathcal{K} \subseteq \mathcal{Q}$.*

Proof. From above, we have $\text{rank}_+(\mathbf{S}) \leq \exp(\Delta^{O(\Delta)} \cdot \sqrt{n} \cdot \log n)$. Moreover, it is well-known that $s \stackrel{\text{def}}{=} \text{rank}_+(\mathbf{S})$ is the extension complexity of some polytope \mathcal{K} such that $\mathcal{P} \subseteq \mathcal{K} \subseteq \mathcal{Q}$. (It is in fact the minimum extension complexity over all such sandwiched polytopes.) The conclusion follows.

For completeness, we show the latter fact: Suppose $\mathbf{S} = \mathbf{U}\mathbf{V}$ is a non-negative factorization of \mathbf{S} with $\mathbf{U} \in \mathbb{R}_{\geq 0}^{f \times s}$ and $\mathbf{V} \in \mathbb{R}_{\geq 0}^{s \times v}$. Let $\mathcal{K}^{\text{lift}} \stackrel{\text{def}}{=} \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+s} : \mathbf{A}\mathbf{x} + \mathbf{U}\mathbf{y} = \mathbf{b}, \mathbf{y} \geq \mathbf{0}\}$, and let \mathcal{K} be the projection of $\mathcal{K}^{\text{lift}}$ onto the first n coordinates. It is immediately clear that $\mathcal{K} \subseteq \mathcal{Q}$, and $\text{xc}(\mathcal{K}) \leq s$ by definition. For each $j \in [v]$, the point \mathbf{x}_j satisfies $\mathbf{A}\mathbf{x}_j + \mathbf{S}^j = \mathbf{b}$, where \mathbf{S}^j is the j -th column of \mathbf{S} given by $\mathbf{U}\mathbf{V}\mathbf{e}_j$. It follows that $(\mathbf{x}_j, \mathbf{V}\mathbf{e}_j) \in \mathcal{K}^{\text{lift}}$, so $\mathbf{x}_j \in \mathcal{K}$. As this holds for each $\mathbf{x}_1, \dots, \mathbf{x}_v$, we conclude $\mathcal{P} \subseteq \mathcal{K}$. \square

We give a direct example in a combinatorial optimization setting: Consider the k -SET-PACKING problem, where we are given a collection of n sets $S_1, \dots, S_n \subseteq [N]$ with $N \gg n$, and want to find a maximum subcollection such that each element $j \in [N]$ is contained in at most k sets. The k -SET-PACKING polytope \mathcal{P} is the convex hull of all feasible subcollections of sets, given by

$$\mathcal{P} = \text{conv} \left\{ x \in \{0, 1\}^n : \sum_{i: j \in S_i} x_i \leq k \quad \forall j \in [N] \right\}.$$

Its natural LP relaxation \mathcal{Q} is

$$\mathcal{Q} = \left\{ x \in [0, 1]^n : \sum_{i: j \in S_i} x_i \leq k \quad \forall j \in [N] \right\}.$$

In the regime where $N \gg n$, a priori, the extension complexity of \mathcal{P} and \mathcal{Q} could be as large as N . But interestingly, let \mathbf{S} be the partial slack matrix with respect to \mathcal{P} and \mathcal{Q} as defined in Lemma 3. Then \mathbf{S} contains integral values in $\{0, \dots, k\}$, and so we conclude there exists a sandwiched polytope $\mathcal{P} \subseteq \mathcal{K} \subseteq \mathcal{Q}$ with $\text{xc}(\mathcal{K}) \leq \exp(k^{O(k)} \cdot \sqrt{n} \cdot \log n)$.

2 Communication protocol

In this section, we give a deterministic communication protocol for a bounded integral matrix \mathbf{S} assuming it has the crucial property that any submatrix contains a large monochromatic rectangle. The protocol is based on the protocol from Nisan and Wigderson [NW95] that is expanded on by Lovett [Lov16].

Lemma 4. *Let $0 < \delta < 1$. Let $\mathbf{M} \in \{0, 1, \dots, \Delta\}^{A \times B}$ be a bounded integral matrix with rank r , and suppose for any submatrix $\mathbf{S} \stackrel{\text{def}}{=} \mathbf{M}[A', B']$ where $A' \subseteq A$ and $B' \subseteq B$, there is a monochromatic rectangle in \mathbf{S} of size $\geq \exp(-\delta(r))|A'||B'|$ for some function δ of r . Then,*

$$CC^{\text{det}}(\mathbf{M}) \leq \Theta \left(\log \Delta + \log^2 r + \sum_{i=0}^{\log r} \delta(r/2^i) \right).$$

We begin by proving two helper lemmas relating to \mathbf{M} and the rank of its submatrices, which we subsequently use to bound the communication complexity.

Lemma 5. *Let $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{R}$ be matrices of the appropriate dimensions, and let \mathbf{R} have rank 0 or 1. Then*

$$\text{rank} \begin{pmatrix} \mathbf{R} \\ \mathbf{B} \end{pmatrix} + \text{rank} \begin{pmatrix} \mathbf{R} & \mathbf{A} \end{pmatrix} \leq \text{rank} \begin{pmatrix} \mathbf{R} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{pmatrix} + 3.$$

Proof. We use a sequence of elementary rank properties:

$$\begin{aligned} \text{rank} \begin{pmatrix} \mathbf{R} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{pmatrix} + 1 &\geq \text{rank} \begin{pmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{pmatrix} \geq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) \\ &\geq \text{rank}(\mathbf{R} \ \mathbf{A}) + \text{rank} \begin{pmatrix} \mathbf{R} \\ \mathbf{B} \end{pmatrix} - 2. \end{aligned}$$

□

Lemma 6. *Suppose the rank of $\mathbf{M} \in \{0, 1, \dots, \Delta\}^{A \times B}$ is r . Then \mathbf{M} contains at most $(\Delta + 1)^r$ different rows and columns.*

Proof. We show the argument for rows: Let $\mathbf{U} \in \{0, \dots, \Delta\}^{A \times r}$ denote a submatrix of \mathbf{M} consisting of r linearly independent columns of \mathbf{M} . Clearly \mathbf{U} has at most $(\Delta + 1)^r$ different rows. If row i and row j of \mathbf{U} are identical, then row i and j of \mathbf{M} are also identical, since by definition of rank, the columns of \mathbf{M} are obtained by taking linear combinations of columns of \mathbf{U} . □

Now, we can design a communication protocol to compute \mathbf{M} using standard techniques. Without loss of generality, we may assume \mathbf{M} does not contain identical rows or columns, so that the conclusion of Lemma 6 can be applied.

of Lemma 4. Suppose Alice has input $a \in A$ and Bob has $b \in B$. Then Alice and Bob compute $\mathbf{M}_{a,b}$ by recursively reducing the matrix \mathbf{M} to a smaller submatrix in one of two ways: they communicate the bit 0 which guarantees a decrease in rank in the resulting submatrix, and the bit 1 which guarantees a decrease in size.

Let \mathbf{S} denote the submatrix to be considered at a recursive iteration. Alice and Bob begin with $\mathbf{S} \stackrel{\text{def}}{=} \mathbf{M}$. During a recursive iteration, they first write \mathbf{S} in the form

$$\mathbf{S} = \begin{pmatrix} \mathbf{R} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{pmatrix},$$

where \mathbf{R} denote the large monochromatic rectangle in \mathbf{S} that is guaranteed to exist, with $|\mathbf{R}| \geq \exp(-\delta(r))|\mathbf{S}|$. By Lemma 5, either $\text{rank}(\mathbf{R} \ \mathbf{A}) \leq \frac{1}{2}\text{rank}(\mathbf{S}) + \frac{3}{2}$, or $\text{rank}(\mathbf{R} \ \mathbf{B}) \leq \frac{1}{2}\text{rank}(\mathbf{S}) + \frac{3}{2}$.

In the first case, Alice communicates. If her input row a is in the upper submatrix of \mathbf{S} , Alice sends the bit 0, and both Alice and Bob update $\mathbf{S} = (\mathbf{R} \ \mathbf{A})$. On the other hand, if row a is in the lower submatrix of \mathbf{S} , Alice sends the bit 1, and they both update $\mathbf{S} = (\mathbf{B} \ \mathbf{C})$. In the second case, Bob communicates. If his input column b is in the left submatrix of \mathbf{S} , Bob sends 0, and they update $\mathbf{S} = \begin{pmatrix} \mathbf{R} \\ \mathbf{B} \end{pmatrix}$. Otherwise, Bob sends 1, and they update $\mathbf{S} = \begin{pmatrix} \mathbf{A} \\ \mathbf{C} \end{pmatrix}$. If \mathbf{S} has size 1, Alice can simply output the entry of \mathbf{S} , which is precisely $\mathbf{M}_{a,b}$.

Consider the communication protocol up until the rank of \mathbf{S} is halved (i.e., when the first 0 bit is communicated): The protocol tree has at most $\Theta\left(\frac{-\log |\mathbf{S}|}{\log(1 - \exp(-\delta(r)))}\right) = \Theta\left(\frac{r \log(\Delta + 1)}{\exp(-\delta(r))}\right)$ -many leaves at this point, which is the max number of 1 bits that could have been communicated. Standard balancing techniques (c.f. [RY20, Chapter 1, Theorem 1.7]) then allow us to balance the protocol tree; combined with the bound $|\mathbf{S}| \leq (\Delta + 1)^{2r}$ by Lemma 6, we conclude that there exists a protocol of length $O(\log r + \log \log(\Delta + 1) + \delta(r))$.

Next, consider the phase where the protocol continues until the rank drops from $r/2$ to $r/4$. Note that we may assume the submatrix at the start of this phase has unique rows and columns, as Alice and Bob has unlimited computation at the start. Then this phase can be simulated by a protocol of length $O(\log(r/2) + \log \log(\Delta + 1) + \delta(r/2))$. We proceed in the same fashion where at

phase i , the rank drops from $r/2^i$ to $r/2^{i+1}$. Summing over $i = 0, \dots, \log r$, we get that the total protocol length is at most

$$\begin{aligned} & \sum_{i=0}^{\log r} \left(\log \left(\frac{r}{2^i} \right) + \log \log(\Delta + 1) + \delta \left(\frac{r}{2^i} \right) \right) \\ & \leq \log^2 r + \log r \log \log(\Delta + 1) + \sum_{i=0}^{\log r} \delta \left(\frac{r}{2^i} \right). \end{aligned}$$

In the base case, \mathbf{S} has rank 1 or 2. Suppose it has rank 2, and let $\mathbf{S} = \mathbf{u}\mathbf{v}^\top + \mathbf{u}'\mathbf{v}'^\top$ be a factorization, where \mathbf{u} and \mathbf{u}' are two linearly independent columns of \mathbf{S} . Then Alice sends \mathbf{u}_a and \mathbf{u}'_a , which Bob uses to compute $\mathbf{u}_a\mathbf{v}_b + \mathbf{u}'_a\mathbf{v}'_b = \mathbf{S}_{a,b} = \mathbf{M}_{a,b}$. Since \mathbf{u} and \mathbf{u}' are columns of \mathbf{S} , their entries take values from $\{0, 1, \dots, \Delta\}$, so Alice communicates $O(\log \Delta)$ bits in total. Finally we can omit the term $\log r \log \log(\Delta + 1)$ as it is dominated by $\log^2 r + \log \Delta$. \square

3 Finding large monochromatic rectangles

In this section, we show that the assumption for applying the communication protocol from Section 2 does indeed hold. That is, any bounded integral matrix contains a sufficiently large monochromatic rectangle.

We first reduce the problem of finding large monochromatic rectangles to finding *almost-monochromatic* rectangles. We say a rectangle is $(1 - \varepsilon)$ -monochromatic if at least a $(1 - \varepsilon)$ -fraction of its entries have the same value.

Lemma 7. *Suppose $\mathbf{M} \in \mathbb{R}^{A \times B}$ has rank $r \geq 1$, and is $(1 - \frac{1}{16r})$ -monochromatic with color α . Then \mathbf{M} contains a monochromatic rectangle of size $\geq \frac{|A||B|}{8}$.*

Proof. Let us call a column of \mathbf{M} *bad* if it contains at least $\frac{1}{8r}|A|$ -many non- α entries. By Markov's inequality, at most half the columns are bad. Let $B' \subseteq B$ be the remaining good columns, where each contains at least $(1 - \frac{1}{8r})|A|$ -many α entries. Let $B'' \subseteq B'$ be a maximal set of linearly independent good columns. We know $|B''| \leq r$ as the rank of \mathbf{M} is r .

Let $\mathbf{U} = \mathbf{M}[A, B'']$. Since each column in B'' contains at most $\frac{1}{8r}|A|$ -many non- α entries, there are at most $r \cdot \frac{1}{8r}|A|$ rows of \mathbf{U} that contain non- α entries. Let A' denote the $|A| - \frac{1}{8}|A| \geq \frac{1}{2}|A|$ rows of \mathbf{U} that contain only α entries. So $\mathbf{M}[A', B'']$ contains only α entries.

Let $\mathbf{T} = \mathbf{M}[A', B']$. Since the columns in B' are linear combinations of columns in B'' , each column of \mathbf{T} must be of the form $\beta \mathbf{1}$ for some β . Finally, we know $|\mathbf{T}| = |A'||B'| \geq \frac{1}{4}|A||B|$, and there are at most $\frac{1}{16r}|A||B|$ non- α entries in \mathbf{M} in total, so at least half the columns of \mathbf{T} must have value α . \square

Now, it remains to show that we can find large almost-monochromatic rectangles.

Lemma 8. *Let $\mathbf{M} \in \{0, 1, \dots, \Delta\}^{A \times B}$ be a bounded integral matrix of rank r . Then \mathbf{M} contains a $(1 - \frac{1}{16r})$ -monochromatic rectangle of size at least $|A||B| \cdot \exp(-\Delta^{O(\Delta)} \cdot \sqrt{r} \cdot \log r)$.*

To prove Lemma 8, we first define a distribution \mathcal{D} over the rectangles of \mathbf{M} , and then use the probabilistic method with respect to \mathcal{D} to show that there exists a large enough almost-monochromatic rectangle. We begin with the technical ingredients:

Definition 9 (Factorization norm). For a matrix $\mathbf{M} \in \mathbb{R}^{A \times B}$, define its γ_2 -norm as

$$\gamma_2(\mathbf{M}) \stackrel{\text{def}}{=} \min\{R \geq 0 : \text{there are families of vectors } \{\mathbf{u}_a\}_{a \in A}, \{\mathbf{v}_b\}_{b \in B} \text{ so that}$$

$$\mathbf{M}_{a,b} = \langle \mathbf{u}_a, \mathbf{v}_b \rangle \text{ and } \|\mathbf{u}_a\|_2 \|\mathbf{v}_b\|_2 \leq R \text{ for all } a, b\}.$$

In other words, $\gamma_2(\mathbf{M})$ gives the Euclidean length needed to factor the matrix \mathbf{M} . Note that by rescaling the \mathbf{u}_a 's and \mathbf{v}_b 's, it is straightforward to guarantee $\|\mathbf{u}_a\|_2 \leq \sqrt{R}$ and $\|\mathbf{v}_b\|_2 \leq \sqrt{R}$ for all a, b in the definition. Here $\mathbf{u}_a, \mathbf{v}_b$ are vectors of any dimension (of course one may choose $\mathbf{u}_a, \mathbf{v}_b$ to have dimension $\text{rank}(\mathbf{M})$). The terminology γ_2 -norm or factorization norm is indeed justified as γ_2 is a norm on the space of real matrices.

The following lemma is well-known:

Lemma 10 (Lemma 4.2, [LMSS07]). Any matrix $\mathbf{M} \in \mathbb{R}^{A \times B}$ satisfies $\gamma_2(\mathbf{M}) \leq \|\mathbf{M}\|_\infty \cdot \sqrt{\text{rank}(\mathbf{M})}$.

It will be convenient for us to factor \mathbf{M} with vectors of the same Euclidean length, which comes at the expense of the dimension:

Lemma 11. For any matrix $\mathbf{M} \in \mathbb{R}^{A \times B}$ and $s \geq \sqrt{\gamma_2(\mathbf{M})}$, there are vectors $\{\mathbf{u}_a\}_{a \in A}, \{\mathbf{v}_b\}_{b \in B}$ such that $\mathbf{M}_{a,b} = \langle \mathbf{u}_a, \mathbf{v}_b \rangle$ and $\|\mathbf{u}_a\|_2 = \|\mathbf{v}_b\|_2 = s$ for all $a \in A$ and $b \in B$.

Proof. Construct vectors $\mathbf{u}_a, \mathbf{v}_b$ each with 2-norm $\|\mathbf{u}_a\|_2, \|\mathbf{v}_b\|_2 \leq s$ and $\mathbf{M}_{a,b} = \langle \mathbf{u}_a, \mathbf{v}_b \rangle$ using Lemma 10. Then add $|A| + |B|$ new coordinates, where each \mathbf{u}_a and \mathbf{v}_b receives a “private” coordinate. Set the private coordinate of \mathbf{u}_a to $\sqrt{s^2 - \|\mathbf{u}_a\|_2^2}$, and similarly for \mathbf{v}_b . \square

We denote $N^n(0, 1)$ as the n -dimensional standard Gaussian. Let $S^{n-1} \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ be the n -dimensional unit sphere. The following argument is usually called hyperplane rounding in the context of approximation algorithms:

Lemma 12 (Sheppard’s formula). Any vectors $\mathbf{u}, \mathbf{v} \in S^{n-1}$ with $\langle \mathbf{u}, \mathbf{v} \rangle = \alpha$ satisfy

$$\Pr_{\mathbf{g} \sim N^n(0,1)} [\langle \mathbf{g}, \mathbf{u} \rangle \geq 0 \text{ and } \langle \mathbf{g}, \mathbf{v} \rangle \geq 0] = h(\alpha) \stackrel{\text{def}}{=} \frac{1}{2} \left(1 - \frac{\arccos(\alpha)}{\pi} \right).$$

More generally, any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ satisfy

$$\Pr_{\mathbf{g} \sim N^n(0,1)} [\langle \mathbf{g}, \mathbf{u} \rangle \geq 0 \text{ and } \langle \mathbf{g}, \mathbf{v} \rangle \geq 0] = h \left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2} \right).$$

We can now define a suitable distribution over the rectangles of \mathbf{M} using the above tools. In particular, we want the probability that a rectangle contains an entry to be a function of the entry value.

Lemma 13. Let $\mathbf{M} \in \{0, 1, \dots, \Delta\}^{A \times B}$ be a bounded integral matrix of rank r . Then for any k , there is a distribution \mathcal{D}_k over the rectangles of \mathbf{M} so that for all $a \in A$ and $b \in B$,

$$\Pr_{\mathbf{R} \sim \mathcal{D}_k} [(a, b) \in \mathbf{R}] = \left(h \left(\frac{\mathbf{M}_{a,b}}{\Delta \sqrt{r}} \right) \right)^k,$$

where h is the function defined in Lemma 12.

Proof. We use Lemma 11 to factor \mathbf{M} , which gives vectors $\mathbf{u}_a, \mathbf{v}_b$ for all $a \in A$ and $b \in B$, such that $\mathbf{M}_{a,b} = \langle \mathbf{u}_a, \mathbf{v}_b \rangle$, and $\|\mathbf{u}_a\|_2 = \|\mathbf{v}_b\|_2 = \Delta^{1/2} r^{1/4}$. Then we sample independent Gaussians $\mathbf{g}_1, \dots, \mathbf{g}_k \sim N^n(0, 1)$ and set

$$\mathbf{R}_i \stackrel{\text{def}}{=} \{a \in A : \langle \mathbf{u}_a, \mathbf{g}_i \rangle \geq 0\} \times \{b \in B : \langle \mathbf{v}_b, \mathbf{g}_i \rangle \geq 0\},$$

and then $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_1 \cap \dots \cap \mathbf{R}_k$. Note that \mathbf{R} is indeed a rectangle. For each i and each $(a, b) \in A \times B$ we have

$$\Pr[(a, b) \in \mathbf{R}_i] = h\left(\frac{\langle \mathbf{u}_a, \mathbf{v}_b \rangle}{\|\mathbf{u}_a\|_2 \|\mathbf{v}_b\|_2}\right) = h\left(\frac{\mathbf{M}_{a,b}}{\Delta \sqrt{r}}\right).$$

The overall expression follows by independence of the k rectangles. \square

Finally, we use the above distribution to show the existence of large almost-monochromatic rectangles.

of Lemma 8. Let $\mathcal{E}_0 \dot{\cup} \dots \dot{\cup} \mathcal{E}_\Delta$ be the partition of the entries $A \times B$ based on entry values, so that $\mathcal{E}_j \stackrel{\text{def}}{=} \{(a, b) \in A \times B : \mathbf{M}_{a,b} = j\}$. Let $m_j \stackrel{\text{def}}{=} (64r\Delta)^{(8\Delta)^j}$ for each $j = 0, \dots, \Delta$, and let i be the index such that $m_i \cdot |\mathcal{E}_i|$ is maximized. From this, we also get

$$|\mathbf{M}| = |A| \cdot |B| = \sum_{j=0}^{\Delta} |\mathcal{E}_j| \leq \sum_{j=0}^{\Delta} \frac{m_i}{m_j} |\mathcal{E}_j| \leq m_i |\mathcal{E}_i|. \quad (3.1)$$

For notational convenience, recall $h(\alpha) \stackrel{\text{def}}{=} \frac{1}{2} \left(1 - \frac{\arccos(\alpha)}{\pi}\right)$, and let $c(j) \stackrel{\text{def}}{=} h\left(\frac{j}{\Delta \sqrt{r}}\right)$. We observe that on $[0, 1]$, the function h is convex, monotone increasing, lowerbounded by $h(0) = 1/4$, upperbounded by $h(1) = 1/2$, and $h'(\alpha) = \frac{1}{2\pi\sqrt{1-\alpha^2}} \geq \frac{1}{2\pi}$. Additionally, the following claim about c will be useful for our calculations later:

Claim 14. $\frac{c(j)}{c(j-1)} \geq 1 + \frac{4}{3\pi\Delta\sqrt{r}}$ for $1 \leq j \leq \Delta$. Also, $\frac{c(\Delta)}{c(0)} \leq 1 + \frac{4}{\pi\sqrt{r}}$.

Proof. We may assume $r \geq 2$. We use first order approximations for c . For the first inequality, we also use the fact that $c(j-1) \leq 3/8$:

$$\frac{c(j)}{c(j-1)} \geq \frac{c(j-1) + c'(j-1)}{c(j-1)} \geq 1 + \frac{1}{2\pi\Delta\sqrt{r} \cdot c(j-1)} \geq 1 + \frac{4}{3\pi\Delta\sqrt{r}}.$$

For the second inequality, we have

$$\frac{c(\Delta)}{c(0)} \leq \frac{c(0) + \Delta \cdot c'(\Delta)}{c(0)} = 1 + \frac{4\Delta}{\Delta\sqrt{r} \cdot 2\pi\sqrt{1-1/r}} \leq 1 + \frac{4}{\pi\sqrt{r}}.$$

\square

Next, let \mathcal{D}_k be the distribution from Lemma 13, and generate $\mathbf{R} \sim \mathcal{D}_k$ for some choice of k to be determined. We will show that there exists a k such that \mathbf{R} is expected to be $(1 - \frac{1}{16r})$ -monochromatic with color i , and is sufficiently large. Specifically, the number of i -entries in \mathbf{R} is greater than the

number all other entries in \mathbf{R} by a factor of $16r$ in expectation, and moreover, this difference is sufficiently large, which in turn means \mathbf{R} is sufficiently large.

$$\begin{aligned}
& \mathbb{E}_{\mathbf{R} \sim \mathcal{D}_k} \left[|\mathcal{E}_i \cap \mathbf{R}| - 16r \sum_{j \neq i} |\mathcal{E}_j \cap \mathbf{R}| \right] \\
&= |\mathcal{E}_i| \cdot c(i)^k - 16r \sum_{j \neq i} c(j)^k |\mathcal{E}_j| \quad (\text{by Lemma 13}) \\
&\geq |\mathcal{E}_i| \cdot c(i)^k \left(1 - 16r \sum_{j \neq i} \frac{c(j)^k m_i}{c(i)^k m_j} \right)
\end{aligned}$$

Suppose $\frac{c(j)^k m_i}{c(i)^k m_j} \leq \frac{1}{64r\Delta}$ for each $j \neq i$, then we can conclude

$$\begin{aligned}
&\geq |\mathcal{E}_i| \cdot 4^{-k} \left(1 - \frac{1}{4} \right) \quad (\text{Since } c(i) \geq \frac{1}{4}) \\
&\geq \frac{|A||B|}{m_i} \frac{1}{2 \cdot 4^k}. \quad (\text{by Eq. (3.1)})
\end{aligned}$$

Claim 15. *There exists a choice of k such that $\frac{c(j)^k m_i}{c(i)^k m_j} \leq \frac{1}{64r\Delta}$ for all $j \neq i$.*

Proof. We consider two cases:

- (1) $0 \leq j < i$: In this case we can bound $\frac{c(j)^k m_i}{c(i)^k m_j} \leq \left(\frac{c(i-1)}{c(i)} \right)^k \frac{m_i}{m_0}$. To get our claim, it suffices to choose k to satisfy

$$\begin{aligned}
&\left(\frac{c(i-1)}{c(i)} \right)^k \frac{m_i}{m_0} \leq \frac{1}{64r\Delta} \\
\Leftrightarrow \quad k &\geq \frac{((8\Delta)^i + 1) \log(64r\Delta)}{\log \frac{c(i)}{c(i-1)}}.
\end{aligned}$$

Using the lower bound from Claim 14, along with $\log(1+x) \geq x/2$ for $x \leq 1$, we conclude it suffices to choose k to satisfy

$$k \geq ((8\Delta)^i + 1) \log(64r\Delta) \cdot \frac{3}{2} \pi \Delta \sqrt{r}. \quad (3.2)$$

- (2) $i < j \leq \Delta$: In this case we can bound $\frac{c(j)^k m_i}{c(i)^k m_j} \leq \left(\frac{c(\Delta)}{c(0)} \right)^k \frac{m_i}{m_{i+1}}$. To get our claim, it suffices to choose k to satisfy

$$\begin{aligned}
&\left(\frac{c(\Delta)}{c(0)} \right)^k \frac{m_i}{m_{i+1}} \leq \frac{1}{64r\Delta} \\
\Leftrightarrow \quad k &\leq \frac{((8\Delta)^{i+1} - (8\Delta)^i - 1) \log(64r\Delta)}{\log \frac{c(\Delta)}{c(0)}}.
\end{aligned}$$

Using the upper bound from Claim 14, we conclude it suffices to choose k to satisfy

$$k \leq ((8\Delta)^{i+1} - (8\Delta)^i - 1) \log(64r\Delta) \cdot \frac{\pi \sqrt{r}}{4}. \quad (3.3)$$

To choose k to simultaneously satisfy the two cases, we first verify that the lower and upper bound for k in Eq. (3.2) and Eq. (3.3) are consistent. Indeed when $i \neq 0$, we have

$$\frac{3}{2}((8\Delta)^i + 1)\Delta \leq \frac{1}{4}((8\Delta)^{i+1} - (8\Delta)^i - 1),$$

so we may choose k to be equal to the lower bound. If $i = 0$, then the lower bound from Eq. (3.2) does not apply, so we choose k to be equal to the upper bound. \square

For any established choice of k , we always have $k \leq (8\Delta)^\Delta \log(64r\Delta)\pi\Delta\sqrt{r}$. Moreover, we have $\log m_i \leq (8\Delta)^\Delta \log(64r\Delta)$. We conclude that

$$\begin{aligned} \mathbb{E}_{\mathbf{R} \sim \mathcal{D}_k} \left[|\mathcal{E}_i \cap \mathbf{R}| - 16r \sum_{j \neq i} |\mathcal{E}_j \cap \mathbf{R}| \right] &\geq \frac{|A||B|}{m_i} \frac{1}{2 \cdot 4^k} \\ &\geq \exp(-\Delta^{O(\Delta)} \cdot \sqrt{r} \cdot \log r) |A||B|, \end{aligned}$$

Then any \mathbf{R} attaining this expectation will simultaneously satisfy $|\mathbf{R}| \geq \exp(-\Delta^{O(\Delta)} \cdot \sqrt{r} \cdot \log r) |A||B|$ and be $(1 - \frac{1}{16r})$ -monochromatic. \square

4 Proof of Theorem 1

By Lemma 8, we know any submatrix \mathbf{S} of \mathbf{M} contains a $(1 - \frac{1}{16r})$ -monochromatic rectangle of size $\exp(-\Delta^{O(\Delta)} \cdot \sqrt{r} \cdot \log r) |\mathbf{S}|$. Therefore, by Lemma 7, \mathbf{S} contains a monochromatic rectangle of size $\frac{1}{8} \exp(-\Delta^{O(\Delta)} \cdot \sqrt{r} \cdot \log r) |\mathbf{S}|$. We substitute $\delta(r) = \Delta^{O(\Delta)} \sqrt{r} \log r$ in Lemma 4 to get

$$CC^{\det}(\mathbf{M}) \leq \Theta \left(\log \Delta + \log^2 r + \sum_{i=0}^{\log r} \delta(r/2^i) \right).$$

The summation simplifies as follows:

$$\begin{aligned} \sum_{i=0}^{\log r} \delta(r/2^i) &= \sum_{i=0}^{\log r} \Delta^{O(\Delta)} \sqrt{r} \cdot 2^{-i/2} \cdot \log \left(\frac{r}{2^i} \right) \\ &\leq \Delta^{O(\Delta)} \sqrt{r} \log r \cdot \sum_{i=0}^{\log r} 2^{-i/2}, \end{aligned}$$

where the sum converges. We ignore the lower order terms in the communication complexity expression to conclude

$$CC^{\det}(\mathbf{M}) \leq \Theta \left(\Delta^{O(\Delta)} \sqrt{r} \log r \right).$$

\square

5 Acknowledgment

We thank Sam Fiorini for helpful discussions and anonymous referees for correcting a small error in Claim 14 in the previous version.

References

- [Bar12] Alexander Barvinok. Approximations of convex bodies by polytopes and by projections of spectrahedra. *arXiv preprint arXiv:1204.0471*, 2012. 2
- [CCZ10] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. Extended formulations in combinatorial optimization. *4OR*, 8(1):1–48, 2010. 2
- [FGP⁺15] Hamza Fawzi, João Gouveia, Pablo A Parrilo, Richard Z Robinson, and Rekha R Thomas. Positive semidefinite rank. *Mathematical Programming*, 153:133–177, 2015. 2
- [FMP⁺15] Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald De Wolf. Exponential lower bounds for polytopes in combinatorial optimization. *Journal of the ACM (JACM)*, 62(2):1–23, 2015. 2
- [LMSS07] Nati Linial, Shazar Mendelson, Gideon Schechtman, and Adi Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27:439–463, 2007. 6
- [Lov16] Shachar Lovett. Communication is bounded by root of rank. *Journal of the ACM (JACM)*, 63(1):1–9, 2016. 1, 3
- [LS88] László Lovász and Michael Saks. Lattices, mobius functions and communications complexity. In *29th Annual Symposium on Foundations of Computer Science*, pages 81–90. IEEE Computer Society, 1988. 1
- [NW95] Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995. 3
- [Rot14] Thomas Rothvoß. A direct proof for lovett’s bound on the communication complexity of low rank matrices. *arXiv preprint arXiv:1409.6366*, 2014. 2
- [Rot17] Thomas Rothvoß. The matching polytope has exponential extension complexity. *Journal of the ACM (JACM)*, 64(6):1–19, 2017. 2
- [RY20] Anup Rao and Amir Yehudayoff. *Communication Complexity and Applications*. Cambridge University Press, 2020. 2, 4
- [Yan88] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 223–228, 1988. 2