A direct proof for Lovett’s bound on the communication complexity of low rank matrices

Thomas Rothvoß*
University of Washington, Seattle

Abstract

The log-rank conjecture in communication complexity suggests that the deterministic communication complexity of any Boolean rank-$r$ function is bounded by polylog($r$). Recently, major progress was made by Lovett who proved that the communication complexity is bounded by $O(\sqrt{r} \cdot \log r)$. Lovett’s proof is based on known estimates on the discrepancy of low-rank matrices. We give a simple, direct proof based on a hyperplane rounding argument that in our opinion sheds more light on the reason why a root factor suffices and what is necessary to improve on this factor.

1 Introduction

In the classical communication complexity setting, we imagine to have two players, Alice and Bob and a function $f : X \times Y \rightarrow \{\pm 1\}$. The players agree on a communication protocol beforehand; then Alice is given an input $x \in X$ and Bob is presented an input $y \in Y$. Then the players can exchange messages to figure out the function value $f(x, y)$ of their common input. The cost of the protocol is the number of exchanged bits for the worst case input. Moreover we denote the cost of the most efficient protocol by $CC_{\text{det}}(f)$.

It is common to view the function $f$ as a matrix $M \in \{\pm 1\}^{X \times Y}$ with entries $M_{xy} = f(x, y)$ — we will interchangeably use the function $f$ and the matrix $M$ and we abbreviate $\text{rank}(f) := \text{rank}(M)$. A monochromatic rectangle for $f$ is a subset $R = X' \times Y'$ with $X' \subseteq X$ and $Y' \subseteq Y$ on which the function is constant. In particular, the leaves of the optimal deterministic protocol tree correspond to a partition of $M$ into $2^{CC_{\text{det}}(f)}$ many monochromatic rectangles. Observe that this partition can be used to write $M$ as the sum of $2^{CC_{\text{det}}(f)}$ many rank-1 matrices, which implies that $CC_{\text{det}}(f) \geq \log \text{rank}(f)$. On the other hand it is also known that $CC_{\text{det}}(f) \leq \text{rank}(f)$. In fact, Lovász and Saks [LS88] even conjectured that the rank lower bound is tight up to a polynomial factor, that means $CC_{\text{det}}(f) \leq (\log \text{rank}(r))^{O(1)}$. The exponent in this log-rank conjecture needs to be at least $\log_3(6) \approx 1.63$ (unpublished by Kushilevitz, cf. [NW95]). Small improvements have been made by Kotlov [Kot97], who showed that $CC_{\text{det}}(f) \leq \log(4/3)\text{rank}(f)$ and Ben-Sasson, Ron-Zewi and Lovett [BLR12] who gave an asymptotic improvement of $CC_{\text{det}}(f) \leq O(\frac{\text{rank}(f)}{\log \text{rank}(f)})$, but had to assume the polynomial Freiman Rusza conjecture.

*Email: rothvoss@uw.edu. Supported by NSF grant 1420180 with title “Limitations of convex relaxations in combinatorial optimization”.

1
In a recent breakthrough, Lovett [Lov14] showed an unconditional bound of $CC_{det}(f) \leq O(\sqrt{r} \log r)$. The key ingredient for his result is a lower bound on the discrepancy of a function/matrix.

**Theorem 1** ([KN97] [LMSS07]). For any rank-r Boolean matrix $M$ and any measure $\mu$ on its entries, there exists a rectangle $R$ so that $|\mu(f^{-1}(1) \cap R) - \mu(f^{-1}(-1) \cap R)| \geq \frac{1}{8\sqrt{r}}$.

Formally, the discrepancy of a function is the minimum such quantity over all possible measures,

$$\text{disc}(f) = \min_{\text{measure } \mu} \max_{\text{rectangle } R} |\mu(f^{-1}(1) \cap R) - \mu(f^{-1}(-1) \cap R)|$$

The discrepancy lower bound is tight in general, that means there are indeed functions $f$ with $\text{disc}(f) \leq O(1/\sqrt{\text{rank}(f)})$, which suggests that using discrepancy as a black box might not be enough for a better bound. Recently Shraibman [Shr14] obtained Lovett’s bound in terms of the corruption lower bound. However, for expressing the bound in terms of the rank of the matrix, the result still relies on the black-box bound on the discrepancy.

In this write-up, we give a direct proof that bypasses the discrepancy lower bound and somewhat makes it more clear, what needs to be done in order to break the $\sqrt{\text{rank}(f)}$ barrier. For more details on communication complexity we refer to the book of Kushilevitz and Nisan [KN97].

## 2 Preliminaries

The main technical result of Lovett [Lov14] in a slightly paraphrased form says that we can always find a large rectangle $R$ that is almost monochromatic.

**Theorem 2.** Given any Boolean function $f : X \times Y \rightarrow \{\pm 1\}$ with rank $r$ and a measure $\mu$ on $X \times Y$ with $\mu(f^{-1}(1)) \geq \delta > 0$. Then there exists a rectangle $R \subseteq X \times Y$ with $\mu(R) \geq 2^{-\Theta(\sqrt{r} \log \frac{1}{\delta})}$ so that $E_{(x,y) \sim R}[f(x,y)] \geq 1 - \delta$.

In particular one can use Theorem 2 to find a rectangle $R$ of size $|R| \geq 2^{\Theta(\sqrt{r} \log r)}|X \times Y|$ which has a $(1 - \frac{1}{8r})$-fraction of 1-entries (assuming for symmetry reasons that at least half of the entries of $f$ were 1). By arguments of Gavinsky and Lovett [GL14] such an almost monochromatic rectangle always contains a sub-rectangle $R' \subseteq R$ with $|R'| \geq \frac{1}{8}|R|$ that is fully monochromatic.

The guarantee of having large monochromatic rectangles in any sub-matrix, can then be turned into a protocol using arguments of Nisan and Wigderson:

**Theorem 3** ([NW95]). Assume that any rank-$r$ function $f : X \times Y \rightarrow \{\pm 1\}$ has a monochromatic rectangle of size $2^{-c(r)}$. Then any Boolean function $g$ has

$$CC_{det}(g) \leq O(\log^2 \text{rank}(g)) + \sum_{i=0}^{\log \text{rank}(g)} O(c(\text{rank}(g)/2^i)),$$
In particular, we will apply Theorem 3 with \( c(r) = \Theta(\sqrt{r \log(r)}) \) and obtain a protocol of cost \( O(\sqrt{r \log(r)}) \) for any rank \( r \) function. We want to emphasize that the whole construction only has a polylog\((r)\) overhead, that means the log-rank conjecture is actually equivalent to being able to find rectangles of size \( 2^{-\text{polylog}(r)} \vert X \times Y \) that have at least a \( 1 - \frac{1}{8r} \) fraction of entries +1 (or \(-1\), resp). The reader can find a more detailed explanation of Theorem 3 in Lovett’s paper [Lov14].

### 3 Proof of the main theorem

In this section, we want to reprove Lovett’s main technical result, Theorem 2. Fix a matrix \( M \in \{\pm 1\}^{X \times Y} \) and denote its rank by \( r \). First, what does it actually mean that the matrix has rank \( r \)? By definition it means that there are \( r \)-dimensional vectors \( u_x, v_y \) for all \( x \in X \) and \( y \in Y \) so that \( \langle u_x, v_y \rangle = M_{xy} \). But what can we actually say about the length of those vectors? To quote Linial, Mendelson, Schechtman and Shraibman, it is “well known to Banach space theorists” that length \( r^{1/4} \) suffices (see [LMSS07, Lemma 4.2]). For the case that the Banach space knowledge of the reader got a bit rusty, we include a more or less self-contained proof. In the exposition we follow closely [FGP+14].

**Lemma 4.** Any rank-\( r \) matrix \( M \in \{\pm 1\}^{X \times Y} \) has a factorization \( M = \langle u_x, v_y \rangle \) so that \( u_x, v_y \in \mathbb{R}^r \) are vectors with \( \|u_x\|_2, \|v_y\|_2 \leq r^{1/4} \) for \( x \in X \), \( y \in Y \).

*Proof.* First of all, by the definition of rank, there are some vectors \( u_x, v_y \in \mathbb{R}^r \) so that \( \langle u_x, v_y \rangle = M_{xy} \) with \( \text{span}\{u_x : x \in X\} = \mathbb{R}^r \) — just that we have no a priori guarantee on their length. Observe that this choice of vectors is far from being unique. For example we could choose any regular matrix \( T \in \mathbb{R}^{r \times r} \) and rescale \( u'_x := Tu_x \) and \( v'_y = (T^{-1})^T v_y \). The inner product would remain invariant as \( \langle u'_x, v'_y \rangle = u'_x T^T (T^{-1})^T v'_y = u_x^T v_y = M_{xy} \).

To find a suitable linear map \( T \), we will make use of John’s Theorem ([Joh48], see also the excellent survey of [Bals97]):

**Theorem 5** (John ’48). For any full-dimensional symmetric convex set \( K \subseteq \mathbb{R}^r \) and any Ellipsoid \( E \subseteq \mathbb{R}^r \) that is centered at the origin, there exists an invertible linear map \( T \) so that \( E \subseteq T(K) \subseteq \sqrt{r}E \).

We want to apply John’s Theorem to \( K = \text{conv}\{\pm u_x : x \in X\} \) (which indeed is a symmetric convex set) and the ellipsoid \( E := r^{-1/4}B \) with \( B := \{x \in \mathbb{R}^r : \|x\|_2 = 1\} \) being the unit ball. First, John’s Theorem provides us with a linear map \( T \) so that \( r^{-1/4}B \subseteq \text{conv}\{\pm Tu_x : x \in X\} \subseteq r^{1/4}B \). Now, we can rescale the vectors by letting \( u'_x := Tu_x \) and \( v'_y := (T^{-1})^T v_y \). For the sake of a simpler notation, let us start all over and assume that the original vectors \( u_x \) and \( v_y \) satisfied \( r^{-1/4}B \subseteq K \subseteq r^{1/4}B \) for \( K = \text{conv}\{\pm u_x : x \in X\} \) from the beginning on.

Then by this assumption we immediately see that \( \|u_x\|_2 \leq r^{1/4} \) and it just remains to argue that also \( \|v_y\|_2 \leq r^{1/4} \) for a fixed \( y \in Y \). To see this, take the vector \( w := \frac{v_y}{\|v_y\|_2} \) and observe that \( w \in r^{-1/4}B \) and hence \( w \in K \). By standard linear optimization reasoning, there must be a vertex \( \pm u_x \) of \( K \) so that \( \langle u_x, v_y \rangle \geq \langle w, v_y \rangle \).
This implies that
\[ r^{-1/4} \|v_y\|_2 = |\langle w, v_y \rangle| \leq |\langle u_x, v_y \rangle| = 1 \]
and the claim is proven. \[\square\]

The hyperplane rounding argument

Eventually we are ready to prove Lovett’s claim. Let \( M \in \{\pm 1\}^{X \times Y} \) be the matrix with rank-\( r \) factorization \( M_{xy} = \langle u_x, v_y \rangle \) so that \( \|u_x\|_2, \|v_y\|_2 \leq r^{1/4} \). We abbreviate \( Q_i = \{(x, y) \in X \times Y : M_{xy} = i\} \) as the \( i \)-entries of the matrix. We assume that we have a measure \( \mu \) with \( \mu(Q_1) \geq \delta \) with \( \delta > 0 \) and we will aim at finding a large rectangle that contains mostly 1-entries. It will be convenient to normalize the vectors to \( \bar{u}_x := \frac{u_x}{\|u_x\|_2} \) and \( \bar{v}_y := \frac{v_y}{\|v_y\|_2} \). We can make the following observation about their inner products:

\[
\langle \bar{u}_x, \bar{v}_y \rangle = \frac{\langle u_x, v_y \rangle}{\|u_x\|_2 \cdot \|v_y\|_2} \begin{cases} \geq \frac{1}{\sqrt{r}} & \text{if } M_{xy} = 1 \\ \leq -\frac{1}{\sqrt{r}} & \text{if } M_{xy} = -1 \end{cases}
\]

In other words, the angle between \( u_x \) and \( v_y \) for a 1-entries \((x, y)\) is a tiny bit smaller than the angle for a \(-1\)-entry. It is a standard argument that has been used many times e.g. in approximation algorithms that if we take a random hyperplane, then the chance that a pair of vectors ends up on the same side, is larger if their angle is smaller. Formally, let \( N^r(0, 1) \) be the distribution of an \( r \)-dimensional Gaussian random variable. Then in a slightly modified form, Sheppard’s Formula tells us:

**Lemma 6.** For any unit vectors \( u, v \in \mathbb{R}^r \) with \( \langle u, v \rangle = \alpha \) we have

\[
\Pr_{g \sim N^r(0, 1)} [\langle g, u \rangle \geq 0 \text{ and } \langle g, v \rangle \geq 0] = \frac{1}{2} \left( 1 - \frac{\arccos(\alpha)}{\pi} \right)
\]

In particular, the quantity \( \frac{1}{2}(1 - \frac{1}{\pi}\arccos(\alpha)) \) is monotonically increasing in \( \alpha \) with \( \frac{1}{2}(1 - \frac{1}{\pi}\arccos(\alpha)) \geq \frac{1}{4} \) for all \( \alpha \geq 0 \) and \( \frac{1}{2}(1 - \frac{1}{\pi}\arccos(\alpha)) \leq \frac{1}{4} - \frac{|\alpha|}{\pi} \) for \( \alpha \leq 0 \).

Next, we want to take \( T := 7 \ln(\frac{2}{\delta}) \cdot \sqrt{r} \) many random hyperplanes and define \( R \) as those vectors \( u_x \) and \( v_y \) that always ended up on the positive side. Formally, we will take independent random Gaussian vectors \( g_1, \ldots, g_T \sim N^r(0, 1) \) and define rectangles

\[ R_t := \{x \in X : \langle \bar{u}_x, g_t \rangle \geq 0\} \times \{y \in Y : \langle \bar{v}_y, g_t \rangle \geq 0\} \]

and \( R := R_1 \cap \ldots \cap R_T \). It remains to argue that in expectation \( R \) satisfies the claim of Theorem 2.
First, using Sheppard’s Formula, we know that for an entry \((x, y) \in Q_1\) one has 
\[
\Pr[(x, y) \in R_t] \geq \frac{1}{4},
\]
while for an entry \((x, y) \in Q_{-1}\) one has 
\[
\Pr[(x, y) \in R_t] \leq \frac{1}{4} - \frac{1}{7\sqrt{T}}.
\]

Since we take the Gaussians independently,
\[
\mathbb{E}[\mu(R \cap Q_1)] \geq \mu(Q_1) \cdot \left(\frac{1}{4}\right)^T \quad \text{and} \quad \mathbb{E}[\mu(R \cap Q_{-1})] \leq \mu(Q_{-1}) \cdot \left(\frac{1}{4} - \frac{1}{7\sqrt{T}}\right)^T.
\]

In particular their ratio behaves like
\[
\frac{\mathbb{E}[\mu(R \cap Q_{-1})]}{\mathbb{E}[\mu(R \cap Q_1)]} \leq \left(\frac{1}{4} - \frac{1}{7\sqrt{T}}\right)^T = \frac{1}{\delta} \cdot \left(1 - \frac{4}{7\sqrt{T}}\right)^T \leq \frac{1}{\delta} \exp\left(-T \cdot \frac{4}{7\sqrt{T}}\right) \leq \frac{\delta}{2}
\]
for our choice of \(T = \frac{7}{\delta} \cdot \sqrt{T} \). On the other hand, \(\mathbb{E}[\mu(R)] \geq \mathbb{E}[\mu(R \cap Q_1)] \geq \delta \cdot \left(\frac{1}{4}\right)^T \geq 2^{-\Theta(\sqrt{T \log \frac{1}{\delta}})} \). We can combine those estimates and consider a single expectation
\[
\mathbb{E}\left[\mu(R \cap Q_1) - \frac{1}{\delta} \cdot \mu(R \cap Q_{-1})\right] \geq 2^{-\Theta(\sqrt{T \log \frac{1}{\delta}})}
\]
We take any \(R\) attaining this, then in particular we must have \(\mu(R) \geq 2^{-\Theta(\sqrt{T \log \frac{1}{\delta}})}\) and \(\mu(R \cap Q_{-1}) \leq \delta \cdot \mu(R)\).

4 Remarks

We want to conclude this paper with a couple of remarks:

- Instead of taking \(T\) random Gaussians, one can also find an almost monochromatic rectangle using a single Gaussian. Sample \(g \sim N^r(0, 1)\) and define
  \[
  R := \{x \in X : \langle \bar{u}_x, g \rangle \geq s\} \times \{y \in Y : \langle \bar{v}_y, g \rangle \geq s\}.
  \]
  where \(s = \Theta(r^{1/4} \sqrt{\log T})\) is a suitable threshold. The rectangle \(R\) will satisfy the same guarantee as before (up to constant factors). Geometrically, this approach might be more intuitive, as it means that one can take all vectors in a random cap of the unit ball.

- The approach can also be used to get the discrepancy lower bound \(\text{disc}(f) \geq \Omega(1/\sqrt{\text{rank}(f)})\) as a corollary. Take any measure \(\mu\) and assume that \(\mu(Q_1) \geq 1/2\). Then sample a single Gaussian \(g \sim N^r(0, 1)\) and let \(R := \{x \in X : \langle g, u_x \rangle \geq 0\} \times \{y \in Y : \langle g, v_y \rangle \geq 0\}\). Then
  \[
  \mathbb{E}[\mu(R \cap Q_1) - \mu(R \cap Q_{-1})] \geq \frac{1}{2} \cdot \frac{1}{4} - \frac{1}{2} \cdot \left(\frac{1}{4} - \frac{1}{7\sqrt{T}}\right) = \frac{1}{14\sqrt{T}}.
  \]

- Note that John’s theorem and also the bounds on \(\|u_x\|_2, \|v_y\|_2\) are tight in general. However, it seems plausible that one can modify the hyperplane rounding in order to improve the bounds, possibly depending on the geometric arrangement of the vectors.
• We do hesitate to call our proof “new”, since all its ingredients have been contained already either in Lovett’s paper [Lov14] or in the paper of Linial et al. [LMSS07]. For example the bound on the factorization norm [LMSS07] is based on John’s theorem; the hyperplane rounding is also used to prove Grothendieck’s inequality that is another ingredient of [LMSS07]. Also [Lov14] used an amplification procedure as we did by sampling $T$ hyperplanes.

• There are indeed rank-$r$ Boolean matrices that have $2^{\Omega(r)}$ many different rows and columns. The construction is due to Lovász and Kotlov [KL96]. An explicit construction is as follows: Take $8r$ disjoint symbols $A \cup A' \cup B \cup B'$ with $|A| = |A'| = |B| = |B'| = 2r$. Then define set families

\[
A := \{ a \subseteq A \cup B : |a \cap A| = r \text{ and } |a \cap B| = 1 \} \cup \{ a \subseteq A' \cup B' : |a \cap A'| = r \text{ and } |a \cap B'| = 1 \}
\]

\[
B := \{ b \subseteq A' \cup B : |b \cap B| = r \text{ and } |b \cap A'| = 1 \} \cup \{ b \subseteq A \cup B' : |b \cap B'| = r \text{ and } |B \cap A| = 1 \}
\]

It is not difficult to check that for all $a \in A$ and $b \in B$ one has $|a \cap b| \in \{0, 1\}$. Moreover for different tuples $a, a' \in A$ there is always a $b \in B$ with $|a \cap b| = 0$ and $|a \cap b'| = 1$ and the reverse is true for different $b, b' \in B$. The matrix $M \in \{\pm 1\}^{A \times B}$ defined by $M_{ab} = \langle (2 \cdot 1_a, -1), (2 \cdot 1_b, 1) \rangle$ has then rank at most $8r + 1$ and $|A| = |B| = 2^{\Theta(r)}$ many different rows and columns.

Acknowledgments. The author is very grateful to Paul Beame, James Lee and Anup Rao for helpful discussions.

References

[Bal97] Keith Ball. An elementary introduction to modern convex geometry. In Flavors of geometry, volume 31 of Math. Sci. Res. Inst. Publ., pages 1–58. Cambridge Univ. Press, Cambridge, 1997.

[BLR12] Eli Ben-Sasson, Shachar Lovett, and Noga Ron-Zewi. An additive combinatorics approach relating rank to communication complexity. In 53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012, pages 177–186, 2012.

[FGP+14] Hamza Fawzi, João Gouveia, Pablo A. Parrilo, Richard Z. Robinson, and Rekha R. Thomas. Positive semidefinite rank. CoRR, abs/1407.4095, 2014.

[GL14] Dmitry Gavinsky and Shachar Lovett. En route to the log-rank conjecture: New reductions and equivalent formulations. In Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I, pages 514–524, 2014.

[Joh48] Fritz John. Extremum problems with inequalities as subsidiary conditions. In Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948, pages 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.
[KL96] Andrew Kotlov and László Lovász. The rank and size of graphs. *Journal of Graph Theory*, 23(2):185–189, 1996.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, 1997.

[Kot97] Andrei Kotlov. Rank and chromatic number of a graph. *Journal of Graph Theory*, 26(1):1–8, 1997.

[LMSS07] Nati Linial, Shahar Mendelson, Gideon Schechtman, and Adi Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27(4):439–463, 2007.

[Lov14] Shachar Lovett. Communication is bounded by root of rank. In *Symposium on Theory of Computing, STOC 2014*, New York, NY, USA, May 31 - June 03, 2014, pages 842–846, 2014.

[LS88] László Lovász and Michael E. Saks. Lattices, möbius functions and communication complexity. In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 81–90, 1988.

[NW95] Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995.

[Shr14] Adi Shraibman. The Corruption Bound, Log Rank, and Communication Complexity. *ArXiv e-prints*, September 2014.