Approximate Nonnegative Rank is Equivalent to the Smooth Rectangle Bound

Gillat Kol* Shay Moran† Amir Shpilka‡ Amir Yehudayoff§

Abstract

We consider two known lower bounds on randomized communication complexity: The smooth rectangle bound and the logarithm of the approximate nonnegative rank. Our main result is that they are the same up to a multiplicative constant and a small additive term.

The logarithm of the nonnegative rank is known to be a nearly tight lower bound on the deterministic communication complexity. Our result indicates that proving the analogue for the randomized case, namely that the log approximate nonnegative rank is a nearly tight bound on randomized communication complexity, would imply the tightness of the information cost bound.

Another corollary of our result is the existence of a boolean function with a quasipolynomial gap between its approximate rank and approximate nonnegative rank.

---

*School of Mathematics, Institute for Advanced Study, Princeton NJ. Research at the IAS supported by The Fund For Math and the Weizmann Institute of Science National Postdoctoral Award Program for Advancing Women in Science. Part of this work was done while the author was visiting the Simons Institute for the Theory of Computing, Berkeley CA, and the Technion, Israel.

†Departments of Computer Science and Mathematics, Technion-IIT, Israel. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 257575.

‡Department of Computer Science, Technion-IIT, Israel. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 257575.

§Department of Mathematics, Technion-IIT, Israel. Horev fellow – supported by the Taub foundation. Research is also supported by ISF and BSF.

1 Introduction

In this work we are mainly interested in understanding two useful techniques that were developed for proving lower bounds on randomized communication complexity: The smooth rectangle bound [JK10] and the approximate nonnegative rank (see Section 1.3 for both definitions). Our main result is that although these two techniques are seemingly different, the lower bounds that may be derived from them are, more or less, equivalent. As a consequence, we are able to apply previous results regarding the smooth rectangle bound to get new results about the approximate nonnegative rank, thus providing information about two of the open problems in [Lee12] (see Corollaries 6 and 7).

We next survey the relevant lower bounds methods for randomized communication complexity. Here and below, f is a boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. We denote by $\mathsf{D}(f)$ the deterministic communication complexity of f, and by $R_\epsilon(f)$ the randomized private coin[1] communication complexity of f with error ϵ. These and other basic definitions can be found in [KN97].

1.1 Randomized Communication Complexity Lower Bounds

1.1.1 Nonnegative Rank

A well-known linear algebraic lower bound on the deterministic communication complexity of f is[2] $\log \mathsf{rank}(M_f)$, where M_f is the $2^n \times 2^n$ boolean matrix given by $M_f(x,y) = f(x,y)$ [MS82]. The long standing log-rank conjecture asserts that this bound is tight, up to a polynomial overhead.

Conjecture 1 (log-rank conjecture, Lovász and Saks [LS88]). *For every function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, it holds that*[3]

$$\mathsf{D}(f) \leq \text{polylog}\left(\mathsf{rank}(M_f)\right).$$

Yannakakis [Yan91] introduced the notion of *nonnegative rank* to communication complexity. We say that a real matrix M is nonnegative if all its entries are nonnegative. The nonnegative rank of a nonnegative real matrix M, denoted $\mathsf{rank}^+(M)$, is the minimum natural number r such that M is the sum of r nonnegative rank-1 matrices.

[1] For simplicity, we only consider private coin protocols. However, all the results carry over to the public coin model via Newman's Theorem [New91].

[2] Here and below rank is over the real numbers.

[3] In this text, logarithms are base two.

The nonnegative rank is clearly at least as large as rank, that is, $\mathsf{rank}^+(M) \geq \mathsf{rank}(M)$. The nonnegative rank can be arbitrarily larger than the rank, if we allow non-boolean matrices. Indeed, for every $k \in \mathbb{N}$ there exists a matrix M such that $\mathsf{rank}(M) = 3$ and $\mathsf{rank}^+(M) \geq k$ (see [BL09]). However, if we restrict our attention to boolean matrices then no such separation between rank and rank^+ is known. The best known separation for boolean matrices is quasipolynomial[4]. This separation follows from the best known separation between the logarithm of the rank and the communication complexity (see discussion following Corollary 7). Moreover, determining the dependency between rank and rank^+ for boolean matrices is equivalent to solving the log-rank conjecture in communication complexity (see Theorem 2).

While we are still far from proving the log-rank conjecture (the best result in this direction [Lov13] is that $\mathsf{D}(f)$ is at most roughly $\sqrt{\mathsf{rank}(M_f)}$), a variant of the conjecture obtained by replacing the rank by the nonnegative rank is known to hold.

Theorem 2 (**log nonnegative rank theorem, Lovász** [Lov90])**.** *For every function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, it holds that*[5]

$$\mathsf{D}(f) \leq \left(\log \mathsf{rank}^+(M_f) + 1\right) \left(\log \mathsf{rank}(M_{1-f}) + 1\right). \tag{1}$$

In particular, $\mathsf{D}(f) \leq O(\log^2 \mathsf{rank}^+(M_f))$.

We study a randomized analogue of Equation (1). In the randomized setting, the notion of nonnegative rank needs to be altered to an approximate one. The ϵ-approximate nonnegative rank of a matrix M, denoted $\mathsf{rank}^+_\epsilon(M)$, is the minimum nonnegative rank of a $2^n \times 2^n$ nonegative matrix M' so that $\|M - M'\|_\infty \leq \epsilon$, i.e., $|M(x,y) - M'(x,y)| \leq \epsilon$ for all $x, y \in \{0,1\}^n$. The ϵ-approximate rank of a matrix M, denoted $\mathsf{rank}_\epsilon(M)$, is defined similarly.

Again, clearly $\mathsf{rank}_\epsilon(M) \leq \mathsf{rank}^+_\epsilon(M)$. One can also prove, using the separation discussed earlier between rank and rank^+, that for every $k \in \mathbb{N}$ there exists a matrix M and $\epsilon > 0$ such that $\mathsf{rank}_\epsilon(M) \leq 3$ and $\mathsf{rank}^+_\epsilon(M) \geq k$. However, not much is known about the relation between rank_ϵ and rank^+_ϵ for boolean matrices.

It was shown by [Kra96] that $\mathsf{R}_\epsilon(f) \geq \log \mathsf{rank}^+_\epsilon(M_f)$. We consider the following conjecture asserting that this bound is nearly tight, as in the deterministic case:

Conjecture 3 (**log approximate nonnegative rank conjecture, see also TH8 in** [**Lee12**])**.** *For every sufficiently small constant $0 < \epsilon < 1$ and every function f :*

[4]There exists a sequence of boolean matrices $\{M_n\}_{n=1}^{n=\infty}$ such that $\log \mathsf{rank}^+(M_i) = \Omega((\log \mathsf{rank}(M_i))^\alpha)$ for some constant $\alpha > 1$.

[5]We use $1 - f$ to denote the boolean function $(1-f)(x,y) = 1 - f(x,y)$.

$\{0,1\}^n \times \{0,1\}^n \to \{0,1\}$,

$$\mathsf{R}_\epsilon(f) \leq \text{polylog}\left(\mathsf{rank}^+_\epsilon(M_f) \cdot \mathsf{rank}^+_\epsilon(M_{1-f})\right).$$

We will later relate this conjecture to an open problem regarding the compression of communication protocols.

1.1.2 The Smooth Rectangle Bound

A different approach for proving lower bounds on randomized communication complexity, which we refer to as the "rectangle based method", is based on bounding from below the weight of the largest (almost) monochromatic combinatorial rectangle.

The smooth rectangle bound, suggested by [JK10], is a rectangle based method shown to be a stronger lower bound than many of the previous methods (for example, the rectangle/corruption bound, the discrepancy bound, and the γ_2 approach [LS07]). Informally speaking, the smooth rectangle bound for a function f with error ϵ, denoted $\mathsf{srec}^1_\epsilon(f)$, considers assignments of weights (nonnegative real values) to combinatorial rectangles (sets of the form $\mathcal{X} \times \mathcal{Y}$ where $\mathcal{X}, \mathcal{Y} \subseteq \{0,1\}^n$) satisfying:

1. For inputs $(x, y) \in f^{-1}(1)$, the total weight assigned to rectangles containing (x, y) is between $1 - \epsilon$ and 1.
2. For inputs $(x, y) \in f^{-1}(0)$, the total weight assigned to rectangles containing (x, y) is at most ϵ.

The value $\mathsf{srec}^1_\epsilon(f)$ is the logarithm of the minimum total weight assigned to all rectangles by any such assignment. A formal definition of the smooth rectangle bound can be found in Section 1.3.

1.1.3 Information Cost

Another recent lower bound method is based on *information cost*. The information cost of a function f with error ϵ, denoted $\mathsf{IC}_\epsilon(f)$, measures the amount of information the players must learn about each other's input while executing any protocol that computes f, with error at most ϵ [CSWY01, BYJKS04, Bra12]. For a formal definition of $\mathsf{IC}_\epsilon(f)$, see e.g. Definition 2.1 in [KLL+12]. The information cost is known to lower bound the randomized public coin communication complexity of f [BR11]. The other direction, namely whether every function with information cost I has a randomized protocol with communication complexity I (a "compressed" protocol), is yet another open problem in communication complexity [Bra12]. We state here a somewhat weaker conjecture than Open Problem 1 of [Bra12]:

Conjecture 4 (**weak compression conjecture**). *For every sufficiently small* $0 < \epsilon < 1$ *and every function* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$,

$$\mathsf{R}_\epsilon(f) \leq \mathrm{poly}\left(\mathsf{IC}_\epsilon(f), \log(n), 1/\epsilon\right).$$

It was recently shown by [KLL+12] that the information cost bound is at least as powerful as almost all the rectangle methods. This was done by showing that the relaxed partition bound is always (roughly) at most the information cost. It is easily seen (by comparing the corresponding linear-programs) that for boolean functions, the relaxed partition bound corresponds to a two-sided smooth rectangle bound, defined as the maximum between the smooth rectangle bound of f and of $1 - f$. In fact, prior to our work, the logarithm of the approximate nonnegative rank was one of the few bounds not known to be weaker than the information cost.

1.2 Our Results

Our main result is the following theorem showing that the smooth rectangle bound is almost equivalent to the logarithm of the approximate nonnegative rank.

Theorem 5 (**main**). *For every* $0 < \epsilon < \frac{1}{10}$ *and a function* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$,

$$\mathsf{srec}^1_{3\epsilon}(f) \leq \log \mathsf{rank}^+_\epsilon(M_f) \leq 2\mathsf{srec}^1_{\epsilon/2}(f) + \log(12n/\epsilon^2).$$

Furthermore, an additive $\log(n/\epsilon)$ *term on the right hand side is needed.*

Theorem 5 is proved in Section 2. Next, we give several corollaries of this theorem.

The log approximate nonnegative rank conjecture and compression. One corollary is that proving the log approximate nonnegative rank conjecture (Conjecture 3) would imply that the information cost bound is nearly tight (Conjecture 4). Formally, we prove the following corollary.

Corollary 6. *There exists a constant* $c > 0$ *such that for every sufficiently small* $0 < \epsilon < 1$,

$$\mathsf{IC}_\epsilon(f) \geq c \cdot \epsilon^2 \left(\log \mathsf{rank}^+_{4\epsilon}(M_f) - \log(3n/8\epsilon^2)\right) - 1.$$

We give the proof of this corollary in Appendix B.

Separating the approximate rank and nonnegative rank. It follows that the approximate nonnegative rank of the negation of the disjointness function on n bit-strings, denoted NDISJ_n, is quasipolynomial in its approximate rank, thus addressing Problem TH9 in [Lee12]: For a small constant $\epsilon > 0$, [Raz02] proved that $\mathsf{rank}_\epsilon(\mathsf{NDISJ}_n) \leq 2^{O(\sqrt{n})}$ (see also the discussion after Conjecture 42 in [LS09b]), while $\mathsf{srec}^1_{3\epsilon}(\mathsf{NDISJ}_n) \geq \Omega(n)$. By Theorem 5, $\mathsf{rank}^+_\epsilon(\mathsf{NDISJ}_n) \geq 2^{\Omega(n)}$.

Corollary 7. *If $0 < \epsilon < 1$ is a sufficiently small constant then for every $n \in \mathbb{N}$,*

$$\log \mathsf{rank}^+_\epsilon(M_{\mathsf{NDISJ}_n}) \geq \Omega\left(\log^2 \mathsf{rank}_\epsilon(M_{\mathsf{NDISJ}_n})\right).$$

We mention that in the non-approximate case, any gap greater than quasipolynomial between the rank and nonnegative rank will disprove the log-rank conjecture as $\mathsf{D}(f) \geq \log \mathsf{rank}^+(M_f)$. The best known gap is only $\mathsf{D}(f) \geq \Omega\left((\log \mathsf{rank}(M_f))^\alpha\right)$ for $\alpha = \log_3(6) < 2$ (Kushilevitz (unpublished), cf. [NW95]).

New upper bound on deterministic complexity. Theorem 2 implies $\mathsf{D}(f) \leq O\left(\log \mathsf{rank}^+(M_f) \cdot \log \mathsf{rank}(M_f)\right)$. By combining Theorem 5 with results from [GL13] and [JK10], we devise a similar bound using the (potentially smaller) approximate nonnegative rank instead of the nonnegative rank. Thus, in order to prove the log-rank conjecture it is enough to show that $\log \mathsf{rank}(M_f) \leq \mathrm{polylog}(\mathsf{rank}^+_\epsilon(M_f) + \mathsf{rank}^+_\epsilon(M_{1-f}))$.

Corollary 8. *For every function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$,*

$$\mathsf{D}(f) \leq O(\log(\mathsf{rank}^+_{1/18}(M_f) + \mathsf{rank}^+_{1/18}(M_{1-f})) \cdot \log^2 \mathsf{rank}(f)).$$

We give a proof of this corollary in Appendix C.

1.2.1 Open Problems

Consider the following stronger version of the log approximate nonnegative rank conjecture (Conjecture 3), asserting that

$$\mathsf{R}_\epsilon(f) \leq \mathrm{polylog}\left(\mathsf{rank}^+_\epsilon(M_f)\right).$$

This conjecture, if true, has two simple corollaries, that describe properties of boolean matrices and are of independent interest. One concerns the behavior of the approximate nonnegative rank when negating f, namely, that $\log \mathsf{rank}^+_\epsilon(M_{1-f})$ is at most polynomial in $\log \mathsf{rank}^+_\epsilon(M_f)$. Observe that the approximate rank satisfies this property as $\mathsf{rank}(M_{1-f}) \leq \mathsf{rank}(M_f) + 1$, and so does the nonnegative rank as $\mathsf{D}(1-f) = \mathsf{D}(f)$ and

$\log \mathsf{rank}^+(M_f) \leq \mathsf{D}(f) \leq O(\log^2 \mathsf{rank}^+(M_f))$. A second corollary concerns error reduction, namely, that we have the bound $\log \mathsf{rank}^+_\epsilon(M_f) \leq (\log \mathsf{rank}^+_{1/3}(M_f))^{O(\log(1/\epsilon))}$. The approximate rank was shown to satisfy this property (it actually satisfies the stronger property $\mathsf{rank}_\epsilon(M_f) \leq (\mathsf{rank}_{1/3}(M_f))^{O(\log(1/\epsilon))}$, see [Alo03, LS09a]). Both of the above corollaries are still open, and one may wish to study either of them prior to the log approximate nonnegative rank conjecture. In Appendix A we show that the method used in [Alo03, LS09a] to prove error reduction for approximate rank cannot work for the approximate nonnegative rank.

1.3 Definitions

We conclude the introduction by giving the needed formal definitions.

Definition 1 (nonnegative rank). *Let $M \in \mathbb{R}^{n \times m}$ be a matrix. M is* nonnegative *if $M(x,y) \geq 0$ for every x,y. M has* rank one *if it is of the form $M = u \otimes v$, where $u \in \mathbb{R}^n, v \in \mathbb{R}^m$ and $\otimes$ denotes tensor product (that is, $M(x,y) = u(x)v(y)$ for every x,y).*

The nonnegative rank *of a nonnegative matrix M is*

$$\mathsf{rank}^+(M) = \min\left\{ r \in \mathbb{N} \ : \ M = \sum_{i=1}^{r} M_i, \quad \forall\ i \quad M_i \text{ is nonnegative and of rank one} \right\}.$$

The ϵ-approximate nonnegative rank *of M is*

$$\mathsf{rank}^+_\epsilon(M) = \min\left\{ \mathsf{rank}^+(M') \ : \ M' \text{ is nonnegative, } \|M - M'\|_\infty \leq \epsilon \right\}.$$

Definition 2 (smooth rectangle bound). *For $0 \leq \epsilon < \frac{1}{2}$ and a function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, the (one)* smooth rectangle bound $\mathsf{srec}^1_\epsilon(f)$ *is the logarithm of the value of the following linear program. Below, R ranges over combinatorial rectangles (sets of the form $\mathcal{X} \times \mathcal{Y}$ where $\mathcal{X}, \mathcal{Y} \subseteq \{0,1\}^n$) and $R(x,y)$ is the indicator for the event $(x,y) \in R$.*

$$\begin{aligned} \min \sum_R w_R \ : \qquad & \forall (x,y) \in f^{-1}(1): \ 1 - \epsilon \leq \sum_R w_R R(x,y) \leq 1, \\ & \forall (x,y) \in f^{-1}(0): \ \sum_R w_R R(x,y) \leq \epsilon, \\ & \forall R: \ w_R \geq 0. \end{aligned}$$

2 Proving the Equivalence, Theorem 5

The "$\mathsf{srec}^1 \leq \log \mathsf{rank}^+$" direction. We start by upper bounding the smooth rectangle bound by the logarithms of the approximate nonnegative rank. Let $r = \mathsf{rank}^+_\epsilon(M_f)$. Let $M'' \in (\mathbb{R}^+)^{2^n \times 2^n}$ be the promised nonnegative matrix satisfying $\|M_f - M''\|_\infty \leq \epsilon$ and $\mathsf{rank}^+(M'') = r$. Observe that the entries of M'' are bounded by $1+\epsilon$. It will be convenient for us to consider the matrix $M' = \frac{1}{1+\epsilon}M''$ whose entries are bounded by 1. Observe that still $\|M_f - M'\|_\infty \leq 1 - \frac{1-\epsilon}{1+\epsilon} \leq 2\epsilon$ and $\mathsf{rank}^+(M') = r$. Let $M_1, \ldots, M_r \in (\mathbb{R}^+)^{2^n \times 2^n}$ be nonnegative rank-1 matrices so that $M' = \sum_{t=1}^r M_t$.

Fix $t \in [r]$ for now. Write $M_t = v \otimes u$ for two nonnegative vectors $v, u \in (\mathbb{R}^+)^{2^n}$. We may assume without loss of generality that

$$\|v\|_\infty, \|u\|_\infty \leq 1, \tag{2}$$

as u, v can always be converted to such vectors for the following reason: Let $a = v(i)$ be the maximum entry in v, and let $b = u(j)$ be the maximum entry in u. Assume without loss of generality that $b \geq a$. It holds that $1 \geq M_t(i,j) = u(i)v(j) = ab$. If $b \leq 1$ we are done. Otherwise, $b > 1$, and we replace v by bv and u by $\frac{1}{b}u$. Observe that now both vectors have entries in the interval $[0,1]$ (as $a \leq 1/b$), and that $(bv) \otimes \left(\frac{1}{b}u\right) = v \otimes u = M_t$.

Let $K = \lceil \frac{2r}{\epsilon} \rceil$. For an integer $1 \leq k \leq K$, define the vector v_k in the following way: For $i \in [2^n]$, set $v_k(i) = 1/K$ if $v(i) \geq k/K$, and $v_k(i) = 0$ if $v(i) < k/K$. Define u_k similarly. Let

$$v' = \sum_{k \in [K]} v_k \quad \text{and} \quad u' = \sum_{k \in K} u_k.$$

It holds that $\|v - v'\|_\infty, \|u - u'\|_\infty \leq 1/K$, as e.g. v' rounds v to the nearest integer multiple of $1/K$ from below. Let

$$M'_t = v' \otimes u' = \left(\sum_{k \in [K]} v_k\right) \otimes \left(\sum_{k' \in K} u_{k'}\right) = \sum_{k,k' \in [K]} v_k \otimes u_{k'}.$$

Using Equation (2),

$$\begin{aligned}
\|M_t - M'_t\|_\infty &\leq \max_{i,j \in [2^n]} \{v(i)u(j) - v'(i)u'(j)\} \\
&\leq \max_{i,j \in [2^n]} \left\{v(i)u(j) - (v(i) - \tfrac{1}{K})(u(j) - \tfrac{1}{K})\right\} \\
&\leq \tfrac{1}{K} \max_{i,j \in [2^n]} \{v(i) + u(j)\} \leq \tfrac{2}{K}.
\end{aligned}$$

Thus, we approximated M_t (with error $2/K$) by a sum of at most K^2 rectangles, each of weight $1/K^2$.

By summing over all $t \in [r]$,

$$\left\| M' - \sum_{t=1}^{r} M'_t \right\|_\infty \leq \frac{2r}{K} \leq \epsilon.$$

Thus,

$$\left\| M_f - \sum_{t=1}^{r} M'_t \right\|_\infty \leq 3\epsilon.$$

We approximated M_f (with error 3ϵ) by a sum of at most $K^2 r$ rectangles, each of weight $1/K^2$. Furthermore, for every (x, y),

$$\sum_{t=1}^{r} M'_t(x, y) \leq \sum_{t=1}^{r} M_t(x, y) = M'(x, y) \leq 1.$$

Thus, the total weight of rectangles containing (x, y) is at most 1. This means that

$$\mathsf{srec}^1_{3\epsilon}(f) \leq \log(K^2 r \cdot 1/K^2) = \log(r).$$

The "$\log \mathsf{rank}^+ \leq \mathsf{srec}^1$" direction. Next we show that the logarithm of the approximate nonnegative rank is not much larger than the smooth rectangle bound. Let W be such that $\mathsf{srec}^1_\epsilon(f) = \log W$, and let (w_R) be weights for rectangles satisfying the conditions of the linear program defining the smooth rectangle bound, for which $\sum_R w_R = W$. Similarly to [LLR12], we consider the probability distribution on rectangles μ, defined by $\mu(R) = w_R/W$ for all R. For every (x, y), let

$$e_{x,y} = \mathop{\mathbf{E}}_{R\sim\mu}[R(x, y)] = \sum_R \frac{w_R}{W} R(x, y),$$

where we recall that $R(x, y)$ is the indicator for $(x, y) \in R$. We get

$$|f(x, y) - W e_{x,y}| = |f(x, y) - \sum_R w_R R(x, y)| \leq \epsilon. \tag{3}$$

In other words, when R is selected according to μ, the value $W \cdot R(x, y)$ is a good estimation for $f(x, y)$.

Let $R_1, \ldots, R_k$ be independent samples from μ for $k = \lceil 2W^2 n/\epsilon^2 \rceil$. For every (x, y), Hoeffding's bound implies that

$$\Pr\left[\left| \frac{1}{k} \sum_{t=1}^{k} R_t(x, y) - e_{x,y} \right| \geq \epsilon/W \right] \leq 2e^{-2\epsilon^2 k/W^2},$$

where the probability is taken over the (independent) choices of $R_1, \ldots, R_k$. By the union bound, since $2e^{-2\epsilon^2 k/W^2} 2^{2n} < 1$, there is a choice of $R_1, \ldots, R_k$ so that for all (x, y),

$$\left| \frac{1}{k} \sum_{t=1}^{k} R_t(x, y) - e_{x,y} \right| < \epsilon/W. \tag{4}$$

Define

$$M' = \frac{W}{k} \sum_{t=1}^{k} R_t.$$

The nonnegative rank of M' is at most k. By Equations (3) and (4), for all (x, y),

$$|f(x, y) - M'(x, y)| \leq |f(x, y) - W e_{x,y}| + W \left| e_{x,y} - \frac{1}{k} \sum_{t=1}^{k} R_t(x, y) \right| \leq 2\epsilon.$$

Therefore, $\|M - M'\|_\infty \leq 2\epsilon$. This means that

$$\log \mathsf{rank}^+_{2\epsilon}(M_f) \leq \log k \leq 2 \log W + \log(3n/\epsilon^2).$$

The additive $\log(n/\epsilon)$ term is needed. The additive $\log(n/\epsilon)$ term on the right hand side of Theorem 5 must be there as the following example shows. Let f be the equality function, that is, M_f is the $2^n \times 2^n$ identity matrix. In [Alo03, Alo09], it was shown that $\mathsf{rank}_\epsilon(M_f) \geq \Omega\left(\frac{n}{\epsilon^2 \log(1/\epsilon)}\right)$. Obviously the same lower bound holds for the ϵ-approximate nonnegative rank of M_f.

We claim that $\mathsf{srec}^1_\epsilon(f)$ is at most $\log(1/\epsilon)$. Let p_R be the distribution on rectangles of the form $R = A \times A$ defined by: Each x is in A with probability ϵ independently of other x's. Let $w_R = p_R/\epsilon$. For every (x, y), if $f(x, y) = 1$ (i.e., $x = y$), then $\mathbf{E}[R(x, y)] = \Pr[x \in A] = \epsilon$ and so $\sum_R w_R R(x, y) = 1$. If $f(x, y) = 0$, then $\mathbf{E}[R(x, y)] = \Pr[x \in A] \Pr[y \in A] = \epsilon^2$ and so $\sum_R w_R R(x, y) = \epsilon$. So w_R is a solution to the above linear program, and the corresponding value is $\sum_R w_R = 1/\epsilon$.

References

[Alo03] Noga Alon. Problems and results in extremal combinatorics–i. *Discrete Mathematics*, 273(1-3):31–53, 2003. 7, 10, 13

[Alo09] Noga Alon. Perturbed identity matrices have high rank: Proof and applications. *Combinatorics, Probability & Computing*, 18(1-2):3–15, 2009. 10

[BL09] LeRoy B. Beasley and Thomas J. Laffey. Real rank versus nonnegative rank. *Linear Algebra and its Applications*, 431(12):2330 – 2335, 2009. Special Issue in honor of Shmuel Friedland. 3

[BR11] Mark Braverman and Anup Rao. Information equals amortized communication. In *FOCS*, pages 748–757, 2011. 4

[Bra12] Mark Braverman. Interactive information complexity. In *STOC*, pages 505–524, 2012. 4, 5

[BYJKS04] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *J. Comput. Syst. Sci.*, 68(4):702–732, 2004. 4

[CSWY01] Amit Chakrabarti, Yaoyun Shi, Anthony Wirth, and Andrew Chi-Chih Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *FOCS*, pages 270–278, 2001. 4

[GL13] Dmitry Gavinsky and Shachar Lovett. En route to the log-rank conjecture: New reductions and equivalent formulations. *Electronic Colloquium on Computational Complexity (ECCC)*, 20:80, 2013. 6, 15

[JK10] Rahul Jain and Hartmut Klauck. The partition bound for classical communication complexity and query complexity. In *IEEE Conference on Computational Complexity*, pages 247–258, 2010. 2, 4, 6

[KLL$^+$12] Iordanis Kerenidis, Sophie Laplante, Virginie Lerays, Jérémie Roland, and David Xiao. Lower bounds on information complexity via zero-communication protocols and applications. In *FOCS*, pages 500–509, 2012. 4, 5, 14

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, 1997. 2

[Kra96] Matthias Krause. Geometric arguments yield better bounds for threshold circuits and distributed computing. *Theoretical Computer Science*, 156(1&2):99–117, 1996. 3

[Lee12] Troy Lee's homepage. Some open problems around nonnegative rank. `http://www.research.rutgers.edu/~troyjlee/open_problems.pdf`, 2012. 2, 3, 6

[LLR12] Sophie Laplante, Virginie Lerays, and Jérémie Roland. Classical and quantum partition bound and detector inefficiency. In *ICALP (1)*, pages 617–628, 2012. 9

[Lov90] László Lovász. Communication complexity: A survey. *Path, flows, and VLSI-networks*, pages 235–265, 1990. 3

[Lov13] Shachar Lovett. Communication is bounded by root of rank. *CoRR*, abs/1306.1877, 2013. 3

[LS88] László Lovász and Michael E. Saks. Lattices, möbius functions and communication complexity. In *FOCS*, pages 81–90, 1988. 2

[LS07] Nati Linial and Adi Shraibman. Lower bounds in communication complexity based on factorization norms. In *STOC*, pages 699–708, 2007. 4

[LS09a] Troy Lee and Adi Shraibman. An approximation algorithm for approximation rank. In *IEEE Conference on Computational Complexity*, pages 351–357, 2009. 7

[LS09b] Troy Lee and Adi Shraibman. Lower bounds in communication complexity. *Foundations and Trends in Theoretical Computer Science*, 3(4):263–398, 2009. 6

[MS82] Kurt Mehlhorn and Erik Meineche Schmidt. Las vegas is better than determinism in vlsi and distributed computing (extended abstract). In *STOC*, pages 330–337, 1982. 2

[New91] Ilan Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991. 2

[NW95] Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995. 6

[Raz02] Alexander A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Science, Mathematics*, 67:2003, 2002. 6

[Yan91] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, 1991. 2

A No Monotone Error Reduction

Part of the argument in [Alo03] concerns error reduction for approximate rank. It is shown that for a boolean function f, if $\mathsf{rank}_{1/3}(M_f) \leq r$ then $\mathsf{rank}_{\epsilon}(M_f) \leq r^{O(\log(1/\epsilon))}$. For the nonnegative case, we may ask an even easier question: Is it true that

$$\mathsf{rank}^+_{\epsilon}(M_f) \leq \left(\mathsf{rank}^+_{1/3}(M_f) + \mathsf{rank}^+_{1/3}(M_{1-f})\right)^{O(\log(1/\epsilon))}? \tag{5}$$

The argument in [Alo03] is based on the following observation: There is a univariate polynomial p of constant degree d, so that for every $b \in \{0,1\}$ and for every x so that $|x-b| < 1/3$ we have $|p(x)-b| < |x-b|/2$. In other words, the polynomial p contracts around zero and around one. The error reduction follows by observing that if we point-wise apply p to the matrix approximating M_f, we get a better approximation of M_f while the rank is increased by at most a power of d.

We show that this method cannot work in the nonnegative case, in the sense that we cannot replace the polynomial p by a nonnegative polynomial. Specifically, Proposition 9 states that a bivariate polynomial p with certain properties does not exist. Before stating the proposition we demonstrate how one could have used such a polynomial p (should it exist) to perform nonnegative error reduction.

Assume for simplicity that f is so that both $\mathsf{rank}^+_{1/4}(M_f)$ and $\mathsf{rank}^+_{1/4}(M_{1-f})$ are 1. Let M'_0 and M'_1 be nonnegative matrices of rank 1 that are (1/4)-close to M_f and M_{1-f}, respectively. Assume that we are given a bivariate polynomial $p : [0,1] \times [0,1] \to \mathbb{R}^+$ with nonnegative coefficients and finite degree so that for every $(x,y) \in [3/4,1] \times [0,1/4]$ we have $|1-p(x,y)| \leq (1-x)/2$, and for every $(x,y) \in (0,1/4] \times [3/4,1]$ we have $p(x,y) < x/2$. It is not hard to verify that M', defined as

$$M'(x,y) = p(M'_0(x,y), M'_1(x,y)),$$

gives a nonnegative (1/8)-approximation for M_f of constant nonnegative rank.

Proposition 9. *There is no bivariate polynomial $p : [0,1] \times [0,1] \to \mathbb{R}^+$ with nonnegative coefficients so that the followings hold: for every $(x,y) \in [3/4,1] \times [0,1/4]$ we have $|p(x,y) - 1| \leq 1 - x$, and for every $(x,y) \in (0,1/4] \times [3/4,1]$ we have $p(x,y) < x$.*

In other words, a nonnegative polynomial that contracts around $x = 1$ must be expanding around $x = 0$. Notice that we do not restrict the degree of p in the proposition above. We mention that on the line $y = 1 - x$, there is a such a polynomial p of degree four with positive coefficients (e.g. the Bernstein approximation of the step function).

Proof of Proposition 9. Assume towards a contradiction that such a polynomial p exists. Write

$$p(x,y) = yg(x,y) + h(x),$$

where g and h are polynomials with nonnegative coefficients. By assumption,

$$\forall\ 3/4 \le x \le 1 :\ |p(x,0) - 1| = |h(x) - 1| \le 1 - x. \tag{6}$$

In addition, for $0 < x \le 1/4$ it holds that $p(x,1) = g(x,1) + h(x) < x$, so

$$\forall\ 0 < x \le 1/4 :\ h(x) < x. \tag{7}$$

We claim that no such h exists. Indeed, by the above $h(1) = 1$ and $h(1/4) < 1/4$. Since h has positive coefficients, it is convex on the ray of positive real numbers, which implies

$$h(3/4) = h(1/3 \cdot 1/4 + 2/3 \cdot 1) \le 1/3 \cdot h(1/4) + 2/3 \cdot h(1) < 1/3 \cdot 1/4 + 2/3 = 3/4.$$

This is a contradiction to that $|h(3/4) - 1| \le 1 - 3/4$. □

B Proof of Corollary 6

Let $\overline{\mathsf{prt}}$ be the relaxed partition bound, as in Definition 3.2 in [KLL+12]. Let $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be any boolean function. By Theorem 1.1 in [KLL+12] it follows that there exists a constant $c' > 0$ such that for every $0 < \epsilon < \frac{1}{2}$ and for any distribution μ,

$$\mathsf{IC}_\mu(f,\epsilon) \ge c'\epsilon^2(\log \overline{\mathsf{prt}}^\mu_{2\epsilon}(f) - 1) - \tfrac{1}{2}.$$

Since $\mathsf{IC}_\epsilon(f) = \max_\mu \{\mathsf{IC}_\mu(f,\epsilon)\}$ and $\overline{\mathsf{prt}}_{2\epsilon}(f) = \max_\mu \{\overline{\mathsf{prt}}^\mu_{2\epsilon}(f)\}$,

$$\mathsf{IC}_\epsilon(f) \ge c'\epsilon^2(\log \overline{\mathsf{prt}}_{2\epsilon}(f) - 1) - \tfrac{1}{2}. \tag{8}$$

By Lemma 3.3 in [KLL+12], it follows that[6]

$$\log \overline{\mathsf{prt}}_{2\epsilon}(f) \ge \mathsf{srec}^1_{2\epsilon}(f). \tag{9}$$

By Theorem 5, for every ϵ such that $4\epsilon < \frac{1}{10}$,

$$\mathsf{srec}^1_{2\epsilon}(f) \ge \frac{1}{2}\left(\log \mathsf{rank}^+_{4\epsilon}(M_f) - \log(3n/4\epsilon^2)\right). \tag{10}$$

[6]Note the difference between our definition for srec and the definition given in [KLL+12]; our srec is the logarithm of their srec.

By combining Inequalities (8), (9) and (10), we get

$$\mathsf{IC}_\epsilon(f) \geq \frac{c'\epsilon^2}{2}\left(\log \mathsf{rank}^+_{4\epsilon}(M_f) - \log(3n/8\epsilon^2)\right) - \tfrac{1}{2}.$$

C Proof of Corollary 8

Let $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be any boolean function. By Theorem 3.1 in [GL13]:

$$\mathsf{D}(f) \leq O(\mathsf{CC}^{\mathsf{zero}}(f) \cdot (\log \mathsf{rank}(f))^2), \tag{11}$$

where the definition of $\mathsf{CC}^{\mathsf{zero}}$ is:

Definition 3 (Zero communication cost, [GL13]). *The zero communication cost of f, denoted $\mathsf{CC}^{\mathsf{zero}}(f)$, is the minimal c such that the following holds. There exists a distribution p on labeled rectangles (R,z) such that for every $(x,y) \in \mathcal{X} \times \mathcal{Y}$,*

1. $\Pr_{(R,z)\sim p}[(x,y) \in R] \geq 2^{-c}$.
2. $\Pr_{(R,z)\sim p}[f(x,y) = z | (x,y) \in R] \geq 2/3$.

By Inequality (11) it suffices to show that

$$\mathsf{CC}^{\mathsf{zero}}(f) \leq \log(\mathsf{rank}^+_{1/18}(M_f) + \mathsf{rank}^+_{1/18}(M_{1-f})) + 1.$$

Let $\epsilon = \frac{1}{6}$. Let $W_1 = 2^{\mathsf{srec}^1_\epsilon(f)}$ and let $W_0 = 2^{\mathsf{srec}^1_\epsilon(1-f)}$. Let $w_{0,R}$ and $w_{1,R}$ be the corresponding weights so that $W_0 = \sum_R w_{0,R}$ and $W_1 = \sum_R w_{1,R}$.

Define a distribution p on labeled rectangles as follows. For every rectangle R and $z \in \{0,1\}$:

$$p(R,z) = \frac{w_{z,R}}{W_0 + W_1}.$$

Observe that p is indeed a distribution on labelled rectangles. By the definition of the smooth rectangle bound, for every $(x,y) \in \mathcal{X} \times \mathcal{Y}$:

1. $\Pr_{(R,z)\sim p}[(x,y) \in R, f(x,y) = z] \geq \frac{1-\epsilon}{W_0+W_1}$
2. $\Pr_{(R,z)\sim p}[(x,y) \in R, f(x,y) = z] \leq \frac{1}{W_0+W_1}$
3. $\Pr_{(R,z)\sim p}[(x,y) \in R, f(x,y) \neq z] \leq \frac{\epsilon}{W_0+W_1}$

For example, for the first item:

$$\Pr_{(R,z)\sim p}[(x,y) \in R, f(x,y) = z] = \frac{\sum_{R:(x,y)\in R} w_{f(x,y),R}}{W_0 + W_1} \geq \frac{1-\epsilon}{W_0 + W_1}.$$

The other two items are proved in a similar way.

We claim that p certifies that

$$\mathsf{CC}^{\mathsf{zero}}(f) \leq \log(W_0 + W_1) + 1. \tag{12}$$

Indeed, consider the first item in the definition of $\mathsf{CC}^{\mathsf{zero}}$. Let $(x, y) \in \mathcal{X} \times \mathcal{Y}$,

$$\Pr_{(R,z)\sim p}[(x, y) \in R] \geq \frac{1-\epsilon}{W_0 + W_1} \geq \frac{1}{2(W_0 + W_1)} = 2^{-\log(W_0+W_1)-1}.$$

For the second item in the definition: For every $(x, y) \in \mathcal{X} \times \mathcal{Y}$,

$$\begin{aligned}
&\Pr_{(R,z)\sim p}[f(x, y) = z | (x, y) \in R] \\
&\quad = \frac{\Pr_{(R,z)\sim p}[(x, y) \in R, f(x, y) = z]}{\Pr_{(R,z)\sim p}[(x, y) \in R]} \\
&\quad = \frac{\Pr_{(R,z)\sim p}[(x, y) \in R, f(x, y) = z]}{\Pr_{(R,z)\sim p}[(x, y) \in R, f(x, y) = z] + \Pr_{(R,z)\sim p}[(x, y) \in R, f(x, y) \neq z]} \\
&\quad \geq \frac{\frac{1-\epsilon}{W_0+W_1}}{\frac{1}{W_0+W_1} + \frac{\epsilon}{W_0+W_1}} && \text{(the above three items)} \\
&\quad \geq \frac{1-\epsilon}{1+\epsilon} \geq 2/3.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathsf{CC}^{\mathsf{zero}}(f) &\leq \log(W_0 + W_1) + 1 \\
&= \log(2^{\mathsf{srec}^1_{1/6}(f)} + 2^{\mathsf{srec}^1_{1/6}(1-f)}) + 1 \\
&\leq \log(\mathsf{rank}^+_{1/18}(M_f) + \mathsf{rank}^+_{1/18}(M_{1-f})) + 1. && \text{(Theorem 5)}
\end{aligned}$$