Fooling Pairs in Randomized Communication Complexity

Shay Moran[1], Makrand Sinha[*2], and Amir Yehudayoff[†3]

1 **Departments of Computer Science, Technion-IIT, Israel, Microsoft Research, Hertzelia, Israel, and Max Planck Institute for Informatics, Saarbrücken, Germany.**
shaymoran1@gmail.com.

2 **Department of Computer Science and Engineering, University of Washington, Seattle, USA.**
makrand@cs.washington.edu.

3 **Department of Mathematics, Technion-IIT, Israel.**
amir.yehudayoff@gmail.com.

Abstract

The fooling pairs method is one of the standard methods for proving lower bounds for deterministic two-player communication complexity. We study fooling pairs in the context of randomized communication complexity. We show that every fooling pair induces far away distributions on transcripts of private-coin protocols. We use the above to conclude that the private-coin randomized ε-error communication complexity of a function f with a fooling set $\mathcal{S}$ is at least order $\log \frac{\log |\mathcal{S}|}{\varepsilon}$. This relationship was earlier known to hold only for constant values of ε.The bound we prove is tight, for example, for the equality and greater-than functions.

As an application, we exhibit the following dichotomy: for every boolean function f and integer n, the $(1/3)$-error public-coin randomized communication complexity of the function $\bigvee_{i=1}^{n} f(x_i, y_i)$ is either at most c or at least n/c, where $c > 0$ is a universal constant.

1 Introduction

Communication complexity provides a mathematical framework for studying communication between two or more parties. It was introduced by Yao [12] and has found numerous applications since. We focus on the two-player case, and provide a brief introduction to it. For more details see the textbook by Kushilevitz and Nisan [8].

In this model, there are two players called Alice and Bob. The players wish to compute a function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, where Alice knows $x \in \mathcal{X}$ and Bob knows $y \in \mathcal{Y}$. To achieve this goal, they need to communicate. The *communication complexity* of f measures the minimum number of bits the players must exchange in order to compute f. The communication is done according to a pre-determined protocol. Protocols may be deterministic or use randomness that is either *public* (known to both players) or *private* (randomness held by one player is not known to the other). In the case of deterministic protocols, we denote by $D(f)$ the minimum communication required to compute f correctly on all inputs. In the case of randomized protocols, we allow the protocol to err with a small probability. We denote by $R_\varepsilon(f)$ and $R_\varepsilon^{\mathsf{pri}}(f)$ the minimum communication required to compute f correctly with public

* Partially supported by BSF.
† Horev fellow – supported by the Taub foundation. Supported by ISF and BSF.


Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

and private-coin protocols with a probability of error at most ε on all inputs. We refer to section 2.1 for formal definitions.

A fundamental problem in this context is proving lower bounds on the communication complexity of a given function f. Lower bounds methods for deterministic communication complexity are based on the fact that any protocol for f defines a partition of $\mathcal{X} \times \mathcal{Y}$ to f-monochromatic rectangles[1]. Thus, a lower bound on the size of a minimal partition of this kind readily translates to a lower bound on the communication complexity of f. Three basic bounds of this type are based on rectangle size, fooling sets, and matrix rank (see [8]). Both matrix rank and rectangle size lower bounds have natural and well-known analogues in the randomized setting: the approximate rank lower bound [9, 7] and the discrepancy lower bound [8] respectively. In this paper we show that fooling sets also have natural counterparts in the randomized setting.

Although public-coin protocols are more general than private-coin ones, Newman [10] proved that for boolean functions every public-coin protocol can be efficiently simulated by a private-coin protocol: If $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ then for every $0 < \varepsilon < 1/2$,

$$R_{2\varepsilon}(f) \leq R^{\mathsf{pri}}_{2\varepsilon}(f) = O\left(R_{\varepsilon}(f) + \log \frac{\log(|\mathcal{X}||\mathcal{Y}|)}{\varepsilon}\right).$$

The additive logarithmic factor on the right-hand-side is often too small to matter, but it does make a difference in the bounds we prove below.

1.1 Fooling pairs and sets

Fooling sets are a well-known tool for proving lower bounds for $D(f)$. A pair $(x,y),(x',y') \in \mathcal{X} \times \mathcal{Y}$ is called *a fooling pair* for $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ if

- $f(x,y) = f(x',y')$, and
- either $f(x',y) \neq f(x,y)$ or $f(x,y') \neq f(x,y)$.

Observe that if (x,y) and (x',y') are a fooling pair then $x \neq x'$ and $y \neq y'$. When $\mathcal{Z} = \{0,1\}$ we distinguish between 0-fooling pairs (for which $f(x,y) = f(x',y') = 0$) and 1-fooling pairs (for which $f(x,y) = f(x',y') = 1$).

It is easy to see that if (x,y) and (x',y') form a fooling pair then there is no f-monochromatic rectangle that contains both of them. An immediate conclusion is the following:

▶ **Lemma 1** ([8]). *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function, let (x,y) and (x',y') be a fooling pair for f and let π be a deterministic protocol for f. Then*

$$\pi(x,y) \neq \pi(x',y').$$

A subset $\mathcal{S} \subseteq \mathcal{X} \times \mathcal{Y}$ is a *fooling set* if every $p \neq p'$ in $\mathcal{S}$ form a fooling pair. Lemma 1 implies the following basic lower bound for deterministic communication complexity.

▶ Theorem 1.1 ([8]). Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function and let $\mathcal{S}$ be a fooling set for f. Then

$$D(f) \geq \log_2(|\mathcal{S}|).$$

[1] $R \subseteq \mathcal{X} \times \mathcal{Y}$ is an f-monochromatic rectangle if $R = A \times B$ for some $A \subseteq \mathcal{X}, B \subseteq \mathcal{Y}$ and f is constant over R.

The same properties do not hold for randomized protocols, but one could expect their natural variants to hold. Let π be an ε-error private-coin protocol for f, and let $(x,y),(x',y')$ be a fooling pair for f. Then, one can expect that the probabilistic analogue of $\pi(x) \neq \pi(x')$ holds, *i.e.* $|\Pi(x,y) - \Pi(x',y')|$ is large, where $|\Pi(x,y) - \Pi(x',y')|$ denotes the statistical distance between the two distributions on transcripts.

Such a statement was previously only known for a specific type of fooling pair (that we call the AND fooling pair in Section 1.2) and was implicit in [2], where it is used as part of a lower bound proof for the randomized communication complexity of the disjointness function. Here, we prove that it holds for an arbitrary fooling pair.

▶ **Lemma 2** (Analogue of Lemma 1)**.** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function, let (x,y) and (x',y') be a fooling pair for f, and let π be an ε-error private-coin protocol for f. Then*

$$|\Pi(x,y) - \Pi(x',y')| \geq 1 - 2\sqrt{\varepsilon}.$$

Lemma 2 is not only an analogue of Lemma 1 but is actually a generalization of it. Indeed, plugging $\varepsilon = 0$ in Lemma 2 implies Lemma 1. Moreover, it implies that the bound from Theorem 1.1 holds also in the 0-error private-coin randomized case.

We use the above to prove an analogue of Theorem 1.1 as well.

▶ Theorem 1.2 (Analogue of Theorem 1.1). Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function and let $\mathcal{S}$ be a fooling set for f. Let $1/|\mathcal{S}| \leq \varepsilon < 1/3$. Then,

$$R^{\mathsf{pri}}_{\varepsilon}(f) = \Omega\left(\log \frac{\log |\mathcal{S}|}{\varepsilon}\right).$$

The lower bound provided by the theorem above seems exponentially weaker than the one in Theorem 1.1, but it is tight. The equality function EQ over n-bit strings has a large fooling set of size 2^n, but it is well-known (see [8]) that

$$R^{\mathsf{pri}}_{\varepsilon}(\mathsf{EQ}) = O\left(\log \frac{n}{\varepsilon}\right).$$

Theorem 1.2 therefore provides a tight lower bound on $R^{\mathsf{pri}}_{\varepsilon}(\mathsf{EQ})$ in terms of both n and ε. It also provides a tight lower bound for the greater-than function. Moreover, Theorem 1.2 is a generalization of Theorem 1.1 and basically implies it by choosing $\varepsilon = 1/|\mathcal{S}|$.

The proof of the lower bound uses a general lower bound on the rank of perturbed identity matrices by Alon [1]. Interestingly, although not every fooling set comes from an identity matrix (e.g. in the greater-than function), there is always some perturbed identity matrix in the background (the one used in the proof of Theorem 1.2).

We remark that for any constant $0 < \varepsilon < 1/3$, a version of Theorem 1.2 has been known for a long time. In particular, Håstad and Wigderson [5] give a proof of the following result[2] which appears in [12] without proof: for every function f with a fooling set $\mathcal{S}$ and for every $0 < \varepsilon < 1/3$,

$$R^{\mathsf{pri}}_{\varepsilon}(f) = \Omega\left(\log\log |\mathcal{S}|\right). \tag{1.1}$$

[2] In fact, the theorem in [12, 5] is more general than the one stated here. We state the theorem in this form since it fits well the focus of this text.

The right-hand side above does not depend on ε. The same lower bound as in (1.1) also directly follows from Theorem 1.1 and from the following general result [8]: for every function f and for every $0 \leq \varepsilon < 1/2$,

$$R_{\varepsilon}^{\mathsf{pri}}(f) = \Omega(\log D(f)).$$

1.2 Two types of fooling pairs

Let $(x, y), (x', y')$ be a fooling pairfor a boolean function f. There are two types of fooling pairs:

- The AND-type for which $f(x', y) \neq f(x, y')$.
- The XOR-type for which $f(x', y) = f(x, y')$.

A partial proof of Lemma 2 is implicit in [2]. The case considered in [2] corresponds to a 0-fooling pair of the AND-type. Let π be a private-coin ε-error protocol for f that is the AND of two bits. In this case, by definition it must hold that $\Pi(0, 0)$ is statistically far away from $\Pi(1, 1)$. The cut-and-paste property (see Corollary 2.2) implies that the same holds for $\Pi(0, 1)$ and $\Pi(1, 0)$, yielding Lemma 2 for the 0-fooling pair of the AND-type – $(0, 1), (1, 0)$.

The case of a pair of the XOR-type was not analyzed before. If π is a private-coin ε-error protocol for XOR of two bits, then it does not immediately follow that $\Pi(0, 0)$ is far away from $\Pi(1, 1)$, nor that $\Pi(0, 1)$ is far away from $\Pi(1, 0)$. Lemma 2 implies that in fact both are true, but the argument can not use the cut-and-paste property. Our argument actually gives a better quantitative result for the XOR function as compared to the AND function.

The importance of the special case of Lemma 2 from [2] is related to proving a lower bound on the randomized communication complexity of the disjointness function DISJ defined over $\{0, 1\}^n \times \{0, 1\}^n$: $\mathsf{DISJ}(x, y) = 1$ if for all $i \in [n]$ it holds that $x_i \wedge y_i = 0$. They reproved that $R_{1/3}(\mathsf{DISJ}) \geq \Omega(n)$. This lower bound is extremely important and useful in many contexts, and was first proved in [6].

On a high level, the proof of [2] can be summarized as follows. Let π be a private-coin protocol with $(1/3)$-error for DISJ. We want to show that $\mathsf{CC}(\pi) = \Omega(n)$. The argument has two different parts: The first part of the argument essentially relates the *internal information cost* (as was later defined in [3]) of computing one copy of the AND function with the communication of the protocol π for DISJ. This is a direct-sum-esque result. More concretely, if μ is a distribution on $\{0, 1\}^2$ such that $\mu(1, 1) = 0$ then

$$\mathsf{IC}_{\mu}(\mathsf{AND}) \leq \frac{\mathsf{CC}(\pi)}{n},$$

where $\mathsf{IC}_{\mu}(\mathsf{AND})$ is the infimum over all $(1/3)$-error private-coin protocols τ for AND of the internal information cost of τ. The second part of the argument shows that if μ is uniform on the set $\{(0, 0), (0, 1), (1, 0)\}$ then $\mathsf{IC}_{\mu}(\mathsf{AND}) > 0$. The challenge in proving the second part stems from the fact that μ is supported on the zeros of AND, so it is trivial to compute AND on inputs from μ. However, the protocols τ in the definition of $\mathsf{IC}_{\mu}(\mathsf{AND})$ are guaranteed to succeed for every x, y and not only on the support of μ. The authors of [2] use the cut-and-paste property (see Corollary 2.2 below) to argue that indeed $\mathsf{IC}_{\mu}(\mathsf{AND}) > 0$.

Here we observe that these arguments can be cast into a more general fooling-pair based method. For example, consider the following function on a pair of n-tuples of elements:

$$f_k(x,y) = \bigvee_{i=1}^{n} \mathsf{EQ}_k(x_i, y_i),$$

where k is a positive integer and $\mathsf{EQ}_k : [k] \times [k] \to \{0,1\}$ denotes the equality function on elements of the set $[k]$.

The direct-sum reduction of [2] also works for the function f_3 and since EQ_3 contains a 0-fooling pair of the AND-type, we can straightaway conclude that the $(1/3)$-error randomized communication complexity and internal information cost of f_3 are $\Omega(n)$. However, for the seemingly similar function f_2, the direct sum reduction described above does not work (and all the fooling pairs are of the XOR-type). In fact, the $(1/3)$-error public-coin randomized communication complexity and internal information cost of f_2 are $O(1)$, since f_2 can be reduced to equality on n-bit strings.

The following theorem shows that this example is part of a general dichotomy. For example, there is no function f for which the randomized communication complexity of $\bigvee_{i=1}^{n} f(x_i, y_i)$ is $\Theta(\sqrt{n})$, when n tends to infinity.

▶ Theorem 1.3. There is a constant $c > 0$ so that for every boolean function f and integer n, the following holds:

1. If f contains a 0-fooling pair of the AND-type then the $(1/3)$-error public-coin randomized communication complexity of $\bigvee_{i=1}^{n} f(x_i, y_i)$ is at least n/c.
2. Else, the $(1/3)$-error public-coin randomized communication complexity of $\bigvee_{i=1}^{n} f(x_i, y_i)$ is at most c.

A dual statement applies to the n-fold AND of f:

▶ Theorem 1.4 (Dual of Theorem 1.3). There is a constant $c > 0$ so that for every boolean function f and integer n, the following holds:

1. If f contains a 1-fooling pair of the AND-type then the $(1/3)$-error public-coin randomized communication complexity of $\bigwedge_{i=1}^{n} f(x_i, y_i)$ is at least n/c.
2. Else, the $(1/3)$-error public-coin randomized communication complexity of $\bigwedge_{i=1}^{n} f(x_i, y_i)$ is at most c.

We provide a proof of Theorem 1.3. Theorem 1.4 can be derived by a similar argument, or alternatively by a reduction to Theorem 1.3 using the relation $\bigwedge_{i=1}^{n} f(x_i, y_i) = \neg \bigvee_{i=1}^{n} \neg f(x_i, y_i)$, which transforms 1-fooling pairs to 0-fooling pairs.

Proof of Theorem 1.3. To prove the first item, note that the sub-matrix correponding to the 0-fooling pair of the AND-type can be mapped to the AND function and then taking the n-fold copy of it corresponds to computing the negation of the disjointness function on n bits. Applying the lower bound of [2] then proves that randomized communication complexity must be $\Omega(n)$.

For the second item, assume f does not contain any 0-fooling pair of the AND-type. Note that this implies that $\bigvee_{i=1}^{n} f(x_i, y_i)$ also does not contain any 0-fooling pair of the AND-type. Indeed, more generally, if f_1 and f_2 do not contain 0-fooling pairs of the AND-type then $f_1(x_1, y_1) \vee f_2(x_2, y_2)$ also does not contain such pairs.

So, it suffices to show that any function g that does not contain 0-fooling pairs of the AND type has public-coin randomized communication complexity $O(1)$. For any such g, the communication matrix of g does not contain a 2×2 sub-matrix with exactly three *zeros*. Without loss of generality, assume that the communication matrix contains no repeated rows or columns. We claim that this matrix contains at most one *zero* in each row and column. This will finish the proof since by permuting the rows and columns, we get the negation of the identity matrix with possibly one additional column of all ones or one additional row of all ones. Therefore, a simple variant of the $O(1)$ public-coin protocol for the equality function will compute g.

To see why there is at most one *zero* in each row and column, assume towards contradiction that it has two *zeros* in some row i, say in the first and second columns. Now, since the first and second *columns* differ, there must be some other row k on which they disagree. This means that the sub-matrix formed by rows i and k and columns 1 and 2 contains exactly three *zeros*, contradicting our assumption. ◀

2 Preliminaries

2.1 Communication Complexity

A *private-coin communication protocol* for computing a function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ is a binary tree with the following generic structure. Each node in the protocol is owned either by Alice or by Bob. For every $x \in \mathcal{X}$, each internal node v owned by Alice is associated with a distribution $P_{v,x}$ on the children of v. Similarly, for every $y \in \mathcal{Y}$, each internal node v owned by Bob is associated with a distribution $P_{v,y}$ on the children of v. The leaves of the protocol are labeled by $\mathcal{Z}$.

On input x, y, a protocol π is executed as follows.

1. Set v to be the root node of the protocol-tree defined above.
2. If v is a leaf, then the protocol outputs the label of the leaf. Otherwise, if Alice owns the node v, she samples a child according to the distribution $P_{v,x}$ and sends a bit to Bob indicating which child was sampled. The case when Bob owns the node is analogous.
3. Set v to be the sampled node and return to the previous step.

A protocol is *deterministic* if for every internal node v, the distribution $P_{v,x}$ or $P_{v,y}$ has support of size one. A *public-coin* protocol is a distribution over private-coin protocols defined as follows: Alice and Bob first sample a shared random r to choose a protocol π_r, and they execute a private protocol π_r as above.

For an input (x, y), we denote by $\pi(x, y)$ the sequence of messages exchanged between the parties. We call $\pi(x, y)$ the *transcript* of the protocol π on input (x, y). Another way to think of $\pi(x, y)$ is as a leaf in the protocol-tree. We denote by $L(\pi(x, y))$ the label of the leaf $\pi(x, y)$ in the tree. The *communication complexity* of a protocol π, denoted by $\mathsf{CC}(\pi)$ is the depth of the protocol-tree of π. For a private-coin protocol π, we denote by $\Pi(x, y)$ the distribution of the transcript of $\pi(x, y)$.

For a function f, the *deterministic* communication complexity of f, denoted by $D(f)$, is the minimum of $\mathsf{CC}(\pi)$ over all deterministic protocols π such that $L(\pi(x, y)) = f(x, y)$ for every x, y. For $\varepsilon > 0$, we denote by $R_\varepsilon(f)$ the minimum of $\mathsf{CC}(\pi)$ over all public-coin

protocols π such that for every (x, y), it holds that $\mathbb{P}[L(\pi(x, y)) \neq f(x, y)] \leq \varepsilon$ where the probability is taken over all coin flips in the protocol π. We call $R_\varepsilon(f)$ the *ε-error public-coin randomized* communication complexity of f. Analogously we define $R^{\mathsf{pri}}_\varepsilon(f)$ as the *ε-error private-coin* randomized communication complexity.

2.2 Rectangle Property

In the case of deterministic protocols, the set of inputs reaching a particular leaf forms a rectangle (a product set inside $\mathcal{X} \times \mathcal{Y}$). In the case of private-coin randomized protocols, the following holds (see for example Lemma 6.7 in [2]).

▶ **Lemma 3** (Rectangle property for private-coin protocols)**.** *Let π be a private-coin protocol over inputs from $\mathcal{X} \times \mathcal{Y}$, and let $\mathcal{L}$ denote the set of leaves of π. There exist functions $\alpha : \mathcal{L} \times \mathcal{X} \to [0, 1]$, $\beta : \mathcal{L} \times \mathcal{Y} \to [0, 1]$ such that for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and every $\ell \in \mathcal{L}$,*

$$\mathbb{P}[\pi(x, y) \text{ reaches } \ell] = \alpha(\ell, x) \cdot \beta(\ell, y).$$

Here too the lemma is in fact a generalization of what happens in the deterministic case where α, β take values in $\{0, 1\}$ rather than in $[0, 1]$.

The next proposition immediately follows from the definitions.

▶ Proposition 2.1. Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function and let (x, y) and (x', y') be such that $f(x, y) \neq f(x', y')$. Then, for any ε-error private-coin protocol π for f,

$$|\Pi(x, y) - \Pi(x', y')| \geq 1 - 2\varepsilon.$$

2.3 Hellinger Distance and Cut-and-paste Property

The *Hellinger* distance between two distributions p, q over a finite set $\mathcal{U}$ is defined as

$$h(p, q) = \sqrt{1 - \sum_{u \in \mathcal{U}} \sqrt{p(u)q(u)}}.$$

Lemma 3 implies the following property of private-coin protocols that is more commonly known as the cut-and-paste property [11, 4].

▶ Corollary 2.2 (Cut-and-paste property). Let (x, y) and (x', y') be inputs to a randomized private-coin protocol π. Then

$$h(\Pi(x, y), \Pi(x', y')) = h(\Pi(x', y), \Pi(x, y')).$$

We also use the following relationship between Statistical and Hellinger Distances.

▶ Proposition 2.3 (Statistical and Hellinger Distances). Let p and q be distributions. Then,

$$h^2(p, q) \leq |p - q| \leq \sqrt{h^2(p, q)(2 - h^2(p, q))}.$$

In particular, if $|p - q| \geq 1 - \varepsilon$ for $0 \leq \varepsilon \leq 1$. Then, $h^2(p, q) \geq 1 - \sqrt{2\varepsilon}$.

2.4 A Geometric Claim

We use the following technical claim that has a geometric flavor. For two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$, we denote by $\langle \mathbf{a}, \mathbf{b} \rangle$ the standard inner product between $\mathbf{a}, \mathbf{b}$. Denote by $\mathbb{R}_+$ the set of non-negative real numbers.

▶ Claim 2.4. Let $\varepsilon_1, \varepsilon_2, \delta_1, \delta_2 > 0$ and let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}_+^m$ be vectors such that

$$\langle \mathbf{a}, \mathbf{b} \rangle \geq 1 - \varepsilon_1, \qquad \langle \mathbf{c}, \mathbf{d} \rangle \geq 1 - \varepsilon_2,$$
$$\langle \mathbf{a}, \mathbf{c} \rangle \leq \delta_1, \qquad \langle \mathbf{b}, \mathbf{d} \rangle \leq \delta_2.$$

Then,

$$\sum_{i \in [m]} |\mathbf{a}(i)\mathbf{b}(i) - \mathbf{c}(i)\mathbf{d}(i)| \geq 2 - (\varepsilon_1 + \varepsilon_2 + \delta_1 + \delta_2).$$

Proof.

$$\begin{aligned}
&\sum_{i \in [m]} |\mathbf{a}(i)\mathbf{b}(i) - \mathbf{c}(i)\mathbf{d}(i)| \\
&\quad \geq \sum_{i \in [m]} \left(\sqrt{\mathbf{a}(i)\mathbf{b}(i)} - \sqrt{\mathbf{c}(i)\mathbf{d}(i)} \right)^2 && (\forall t, s \geq 0 \;\; |t - s| \geq \left(\sqrt{t} - \sqrt{s}\right)^2) \\
&\quad = \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{d} \rangle - \sum_{i \in [m]} 2\sqrt{\mathbf{a}(i)\mathbf{b}(i)\mathbf{c}(i)\mathbf{d}(i)} \\
&\quad = \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{d} \rangle - \sum_{i \in [m]} 2\sqrt{\mathbf{a}(i)\mathbf{c}(i) \cdot \mathbf{b}(i)\mathbf{d}(i)} \\
&\quad \geq \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{d} \rangle - \sum_{i \in [m]} (\mathbf{a}(i)\mathbf{c}(i) + \mathbf{b}(i)\mathbf{d}(i)) && \text{(AM-GM inequality)} \\
&\quad = \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{d} \rangle - (\langle \mathbf{a}, \mathbf{c} \rangle + \langle \mathbf{b}, \mathbf{d} \rangle) \\
&\quad \geq 2 - (\varepsilon_1 + \varepsilon_2 + \delta_1 + \delta_2).
\end{aligned}$$

◀

3 Fooling pairs and sets

3.1 Fooling pairs induce far away distributions

Proof of Lemma 2. Let the fooling pair be (x, y) and (x', y') and assume without loss of generality that $f(x, y) = f(x', y') = 1$. We distinguish between the following two cases.

(a) $f(x', y) \neq f(x, y')$.
(b) $f(x', y) = f(x, y') = z$ where $z \neq 1$.

In the first case, Corollary 2.1 implies that $|\Pi(x', y) - \Pi(x, y')| \geq 1 - 2\varepsilon$. Proposition 2.2 implies that $h(\Pi(x, y), \Pi(x', y')) = h(\Pi(x', y), \Pi(x, y'))$. Proposition 2.3 thus implies that $|\Pi(x, y) - \Pi(x', y')| \geq 1 - 2\sqrt{\varepsilon}$.

Let us now consider the second case. Let $\mathcal{L}$ be the set of all leaves of π and let $\mathcal{L}_1$ denote those leaves which are labeled by 1. For $x \in \mathcal{X}$, $y \in \mathcal{Y}$, define the vectors $\mathbf{a}_x \in \mathbb{R}_+^{\mathcal{L}_1}$

as $\mathbf{a}_x(\ell) = \alpha(\ell, x)$, and the vectors $\mathbf{b}_y \in \mathbb{R}_+^{\mathcal{L}_1}$ as $\mathbf{b}_y(\ell) = \beta(\ell, y)$ where α and β are the functions from Lemma 3. Since $f(x, y) = 1$ and π is an ε-error protocol for f,

$$\langle \mathbf{a}_x, \mathbf{b}_y \rangle = \sum_{\ell \in \mathcal{L}_1} \alpha(\ell, x) \cdot \beta(\ell, y) = \mathbb{P}[L(\pi(x, y)) = 1] \geq 1 - \epsilon.$$

Similarly, we have $\langle \mathbf{a}_{x'}, \mathbf{b}_{y'} \rangle \geq 1 - \varepsilon$, $\langle \mathbf{a}_x, \mathbf{b}_{y'} \rangle \leq \varepsilon$ and $\langle \mathbf{a}_{x'}, \mathbf{b}_y \rangle \leq \varepsilon$. Observe

$$2|\Pi(x, y) - \Pi(x', y')| \geq \sum_{\ell \in \mathcal{L}_1} |\mathbf{a}_x(\ell)\mathbf{b}_y(\ell) - \mathbf{a}_{x'}(\ell)\mathbf{b}_{y'}(\ell)|.$$

Applying Claim 2.4 with the vectors $\mathbf{a}_x, \mathbf{b}_y, \mathbf{a}_{x'}, \mathbf{b}_{y'}$ yields that $|\Pi(x, y) - \Pi(x', y')| \geq 1 - 2\varepsilon$. ◀

3.2 A lower bound based on fooling sets

The following result of Alon [1] on the rank of perturbed identity matrices is a key ingredient.

▶ **Lemma 4.** *Let $\frac{1}{2\sqrt{m}} \leq \varepsilon < \frac{1}{4}$. Let M be an $m \times m$ matrix such that $|M(i, j)| \leq \varepsilon$ for all $i \neq j$ in $[m]$ and $|M(i, i)| \geq \frac{1}{2}$ for all $i \in [m]$. Then,*

$$\mathsf{rank}(M) = \Omega\left(\frac{\log m}{\varepsilon^2 \log(\frac{1}{\varepsilon})}\right).$$

Proof of Theorem 1.2. Let $\mathcal{L}$ denote the set of leaves of π. Let $A \in \mathbb{R}^{\mathcal{S} \times \mathcal{L}}$ be the matrix defined by

$$A_{(x,y),\ell} = \sqrt{\mathbb{P}[\pi(x, y) = \ell]}.$$

Let

$$M = AA^T$$

where A^T is A transposed. First,

$$M_{(x,y),(x,y)} = 1.$$

Second, if $(x, y) \neq (x', y')$ in $\mathcal{S}$ then by Lemma 2 we know $|\Pi(x, y) - \Pi(x', y')| \geq 1 - 2\sqrt{\varepsilon}$ so by Lemma 2.3

$$h^2(\Pi(x, y), \Pi(x', y')) \geq 1 - 2\varepsilon^{1/4}$$

which implies

$$M_{(x,y),(x',y')} = 1 - h^2(\Pi(x, y), \Pi(x', y')) \leq 2\varepsilon^{1/4}.$$

Lemma 4 implies that[3] the rank of M is at least $\Omega\left(\frac{\log|\mathcal{S}|}{\sqrt{\varepsilon}\log\left(\frac{1}{\varepsilon^{1/4}}\right)}\right) = \Omega\left(\left(\frac{\log|\mathcal{S}|}{\varepsilon}\right)^{1/4}\right)$. On the other hand,

$$2^{CC(\pi)} \geq |\mathcal{L}| \geq \mathsf{rank}(M).$$

◀

[3] We may assume that say $\varepsilon < 2^{-12}$ by repeating the given randomized protocol a constant number of times.

References

1 Noga Alon. Perturbed identity matrices have high rank: Proof and applications. *Comb. Probab. Comput.*, 18(1-2):3–15, March 2009.

2 Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An Information Statistics Approach to Data Stream and Communication Complexity. In *FOCS*, pages 209–218, 2002.

3 Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. *SIAM J. Comput.*, 42(3):1327–1363, 2013.

4 Benny Chor and Eyal Kushilevitz. A zero-one law for boolean privacy. *SIAM J. Discrete Math.*, 4(1):36–47, 1991.

5 J. Hastad and A. Wigderson. The randomized communication complexity of set disjointness. *Theory of Computing*, 3(1):211–219, 2007.

6 Bala Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math.*, 5(4):545–557, 1992.

7 Matthias Krause. Geometric Arguments Yield Better Bounds for Threshold Circuits and Distributed Computing. *Theor. Comput. Sci.*, 156(1&2):99–117, 1996.

8 Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, New York, NY, USA, 1997.

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

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

11 Ramamohan Paturi and Janos Simon. Probabilistic communication complexity. *J. Comput. Syst. Sci.*, 33(1):106–123, 1986.

12 Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *STOC*, pages 209–213, 1979.