


# From Expanders to Hitting Distributions and Simulation Theorems

Alexander Kozachinskiy<sup>1</sup>

National Research University Higher School of Economics, Moscow, Russia

Moscow, 3 Kochnovsky Proezd, Russia

akozachinskiy@hse.ru

 <https://orcid.org/0000-0002-9956-9023>

---

## Abstract

In this paper we explore hitting distributions, a notion that arose recently in the context of deterministic “query-to-communication” simulation theorems. We show that any expander in which any two distinct vertices have at most one common neighbor can be transformed into a gadget possessing good hitting distributions. We demonstrate that this result is applicable to affine plane expanders and to Lubotzky-Phillips-Sarnak construction of Ramanujan graphs. In particular, from affine plane expanders we extract a gadget achieving the best known trade-off between the arity of outer function and the size of gadget. More specifically, when this gadget has  $k$  bits on input, it admits a simulation theorem for all outer function of arity roughly  $2^{k/2}$  or less (the same was also known for  $k$ -bit Inner Product). In addition we show that, unlike Inner Product, underlying hitting distributions in our new gadget are “polynomial-time listable” in the sense that their supports can be written down in time  $2^{O(k)}$ , i.e. in time polynomial in size of gadget’s matrix.

We also obtain two results showing that with current technique no better trade-off between the arity of outer function and the size of gadget can be achieved. Namely, we observe that no gadget can have hitting distributions with significantly better parameters than Inner Product or our new affine plane gadget. We also show that Thickness Lemma, a place which causes restrictions on the arity of outer functions in proofs of simulation theorems, is unimprovable.

**2012 ACM Subject Classification** Theory of computation → Communication complexity

**Keywords and phrases** simulation theorems, hitting distributions, expanders

**Digital Object Identifier** 10.4230/LIPIcs.MFCS.2018.4

## 1 Introduction

Assume that we have a Boolean function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  called *outer function* and a Boolean function  $g : A \times B \rightarrow \{0, 1\}$  called *gadget*. Consider a composed function  $f \circ g : A^n \times B^n \rightarrow \{0, 1\}$ , defined as follows:

$$(f \circ g)((a_1, \dots, a_n), (b_1, \dots, b_n)) = f(g(a_1, b_1), \dots, g(a_n, b_n)).$$

How can we deal with deterministic communication complexity of  $f \circ g$ , denoted below by  $D^{cc}(f \circ g)$ ? Obviously, we have the following inequality:

$$D^{cc}(f \circ g) \leq D^{dt}(f) \cdot D^{cc}(g),$$

where  $D^{dt}(f)$  stands for deterministic query complexity of  $f$ . Indeed, we can transform a decision tree for  $f$  making  $q$  queries into a protocol of communication cost  $q \cdot D^{cc}(g)$  by

---

<sup>1</sup> Supported in part by RFBR grant 16-01-00362 and by the Russian Academic Excellence Project “5-100”.



© Alexander Kozachinskiy;

licensed under Creative Commons License CC-BY

43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018).

Editors: Igor Potapov, Paul Spirakis, and James Worrell; Article No. 4; pp. 4:1–4:15

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

simulating each query to  $f$  with  $D^{cc}(g)$  bits. It turns out that for some gadgets  $g$  and for all  $f$  of arity at most some function of  $g$ 's size this simple protocol is essentially optimal. The first gadget for which this was proved is the Indexing Function

$$\text{IND}_k : \{1, 2, \dots, k\} \times \{0, 1\}^k \rightarrow \{0, 1\}, \quad g(x, y) = y_x.$$

More specifically, in 2015 Göös et al. ([4]) proved that for all  $n \leq 2^{k^{1/20}}$  and for all  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  it holds that

$$D^{cc}(f \circ \text{IND}_k) = \Omega(D^{dt}(f) \log k). \tag{1}$$

Actually, instead of  $f$  we can have not only a Boolean function but any relation  $R \subset \{0, 1\}^n \times C$ . The work of Göös et al. was a generalization of the theorem of Raz and McKenzie ([9]), who in 1997 established (1) for a certain class of outer relations, called DNF-Search problems.

Theorems of this kind, called usually *simulation theorems*, can be viewed as a new method of proving lower bounds in communication complexity. Namely, lower bound on communication complexity of a composed function reduces to lower bound on query complexity of an outer function, and usually it is much easier to deal with the latter. As was shown by Raz and McKenzie, this method turns out to be powerful enough to separate monotone NC-hierarchy. Moreover, as was discovered by Göös et al., this method can be quadratically better than the logarithm of the partition number, another classical lower bound method in deterministic communication complexity.

There are simulation theorems not only for deterministic communication and query complexities, but for other models too, see, e.g., [2, 5, 6, 3].

Note that input length of a gadget in (1) is even bigger than input length of an outer function. Göös et al. in [4] asked, whether it is possible to prove a simulation theorem for a gadget which input length is logarithmic in input length of an outer function. This question was answered positively by Chattopadhyay et al. ([1]) and independently by Wu et al. ([12]). Moreover, Chattopadhyay et al. significantly generalized the proof of Göös et al., having discovered a certain property of a gadget  $g : A \times B \rightarrow \{0, 1\}$  which can be used as a black-box to show new simulation theorems: once  $g$  satisfies this property, we have a simulation theorem for  $g$ . This property can be defined as follows. Let  $\mu$  be probability distribution over rectangles  $U \times V \subset A \times B$ . Distribution  $\mu$  is called  $(\delta, h)$ -*hitting*, where  $\delta \in (0, 1)$  and  $h$  is a positive integer, if for every  $X \subset A$  of size at least  $2^{-h}|A|$  and for every  $Y \subset B$  of size at least  $2^{-h}|B|$  we have that

$$\Pr_{U \times V \sim \mu} [U \times V \cap X \times Y \neq \emptyset] \geq 1 - \delta.$$

It turns out that if for every  $b \in \{0, 1\}$  there is  $(\delta, h)$ -hitting distribution over  $b$ -monochromatic rectangles of  $g$ , then there is a simulation theorem for  $g$ . The smaller  $\delta$  and the bigger  $h$ , the better simulation theorem. More precisely, Chattopadhyay et al. proved the following theorem.

► **Theorem 1.** *Assume that  $\varepsilon \in (0, 1)$  and an integer  $h$  are such that  $h \geq 6/\varepsilon$ . Then the following holds. For every (possibly partial) Boolean function  $g : A \times B \rightarrow \{0, 1\}$  that has two  $(\frac{1}{10}, h)$ -hitting distribution, the one over 0-monochromatic rectangles and the other over 1-monochromatic rectangles, for every  $n \leq 2^{h(1-\varepsilon)}$  and  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  it holds that*

$$D^{cc}(f \circ g^n) \geq \frac{\varepsilon h}{4} \cdot D^{dt}(f).$$

Further, they showed that Inner Product and Gap Hamming Distance gadgets on  $k$  bits have  $(o(1), \Omega(k))$ -hitting distributions for both kinds of monochromatic rectangles. More precisely, for every constant  $\gamma > 0$  and for all large enough  $k$  they constructed  $(o(1), (1/2 - \gamma)k)$ -hitting distributions for  $k$ -bit Inner Product (denoted below by  $\text{IP}_k$ ). Due to Theorem 1 this yields the following simulation theorem for  $\text{IP}_k$ : for every constant  $\gamma > 0$  and for all  $k$  large enough

$$D^{cc}(f \circ \text{IP}_k) = \Omega(D^{dt}(f) \cdot k),$$

where  $f$  is any Boolean function depending on at most  $2^{(1/2-\gamma)k}$  variables. Other gadgets studied until this work do not achieve the same trade-off between the size of outer functions and the size of gadget. Namely, for  $k$ -bit Gap Hamming Distance the lower bound  $D^{cc}(f \circ \text{GHD}) = \Omega(D^{dt}(f) \cdot k)$  is shown in [1] only for  $f$  depending on roughly  $2^{0.45k}$  variables or less. For Indexing gadget, as we saw, this trade-off is exponentially worse.

We also touch upon the following question. Theorem 1 transforms a communication protocol for  $f \circ g^n$  into decision tree for  $f$  (here  $n$  is the arity of  $f$ ). It is an interesting open question whether resulting decision tree can be made efficient, provided that the initial protocol is efficient. More precisely, a decision tree should decide which variable to query in time polynomial in  $n$  (provided that messages of the initial protocol can be computed by players in time polynomial in  $n$ ). Let us note here that we are mostly interested in the regime when  $n$  is exponential in  $k$  (the size of input to  $g$ ) and hence in this regime the size of  $g$ 's matrix is polynomial in  $n$ .

Unfortunately, known constructions does not provide this – resulting decision trees work in exponential time. Among other obstacles to resolve this issue there is a particular step in transformation which deals with hitting distributions. Namely, a tree constantly runs a subroutine which, having a family of subrectangles (not necessarily monochromatic) of  $g$ 's matrix on input, outputs single monochromatic rectangle which intersects most of them. Namely, a subroutine samples rectangle at random from a hitting distribution, which with good probability gives us a correct answer. To derandomize this procedure instead of sampling a rectangle we can just try all the rectangles from the support of hitting distribution. If the support is of size  $2^{O(k)}$  (i.e. polynomial in size of  $g$ 's matrix) and, moreover, the support can be listed in time  $2^{O(k)}$ , we call a corresponding hitting distribution (or, more precisely, family of distributions) *polynomial-time listable*.

For example, hitting distribution from [1] for  $k$ -bit Gap Hamming Distance gadget are polynomial-time listable (roughly speaking, we just have to list all Hamming balls of a certain radius). At the same time, hitting distributions for  $k$ -bit Inner Product from [1] are not polynomial-time listable. Namely, their supports are of size  $2^{\Omega(k^2)}$  (this number corresponds to the number of  $k/2$ -dimensional subspaces of  $\mathbb{F}_2^k$ ). Though due to Chernoff bound it is possible to transform any  $(0.1, h)$ -hitting distribution into, say,  $(0.2, h)$ -hitting distribution with support size  $2^{O(k)}$  (see Proposition 7 below), this does not give explicit construction.

## 1.1 Our results

We show how to transform any explicit expander satisfying one additional restriction into a gadget with polynomial-time listable hitting distributions. The transformation is as follows. Assume that we have a graph  $G = (V, E)$  and a coloring  $c : V \rightarrow \{0, 1\}$ . For  $v \in V$  let  $\Gamma(v)$  denote the set of all  $u \in V$  such that  $u$  and  $v$  are connected by an edge in  $G$ . Assume further that for any two distinct  $u, v \in V$  it holds that  $|\Gamma(u) \cap \Gamma(v)| \leq 1$ . Then the following partial

function is well defined:

$$g(G, c) : V \times V \rightarrow \{0, 1\},$$

$$g(G, c)(u, v) = \begin{cases} 1 & u \neq v \text{ and there is } w \in \Gamma(u) \cap \Gamma(v) \text{ s.t. } c(w) = 1, \\ 0 & u \neq v \text{ and there is } w \in \Gamma(u) \cap \Gamma(v) \text{ s.t. } c(w) = 0, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Call  $c$  balanced if each color is used at least  $|V|/3$  times in  $c$ . It turns out that if  $G$  is a good expander and if  $c$  is balanced, then  $g(G, c)$  possesses good hitting distributions:

► **Theorem 2.** *Assume that  $G = (V, E)$  is a  $(m, d, \gamma)$ -spectral expander in which for any two distinct  $u, v \in V$  it holds that  $|\Gamma(u) \cap \Gamma(v)| \leq 1$  and  $c : V \rightarrow \{0, 1\}$  is a balanced coloring of  $G$ . Assume also that  $m \geq 1/\gamma^2$ . Then for any  $b \in \{0, 1\}$  there is a  $(\frac{1}{10}, \lfloor 2 \log_2(1/\gamma) \rfloor - 100)$ -hitting distribution  $\mu_b$  over  $b$ -monochromatic rectangles of  $g(G, c)$ . All the probabilities of  $\mu_b$  are rational. Moreover, there is a deterministic Turing machine which, having  $b, G$  and  $c$  on input, in time  $m^{O(1)}$  lists all the rectangles from the support of  $\mu_b$ , together with probabilities  $\mu_b$  assigns to them.*

Provided that  $G$ 's adjacency matrix and  $c$ 's truth table can be computed in time  $m^{O(1)}$ , from Theorem 2 we obtain polynomial-time listable family of hitting distributions.

In particular, we apply Theorem 2 to the following explicit family of expanders. If  $q$  is a power of prime, let  $AP_q$  denote a graph in which vertices are pairs of elements of  $\mathbb{F}_q$  and in which  $(a, b), (x, y) \in \mathbb{F}_q^2$  are connected by an edge if and only if  $ax = b + y$ . It is known that  $AP_q$  is a  $(q^2, q, 1/\sqrt{q})$ -spectral expander. It can be easily shown that for any two distinct vertices  $u, v$  of  $AP_q$  it holds that  $|\Gamma(u) \cap \Gamma(v)| \leq 1$ .

► **Corollary 3.** *Let  $q$  be a power of prime. Then in  $AP_q$  for any two distinct vertices  $u, v$  it holds that  $|\Gamma(u) \cap \Gamma(v)| \leq 1$ . Moreover, for all  $n \leq 2^{\log_2 q - 200}$  and  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  the following holds: if  $c$  is a balanced coloring of  $AP_q$ , then*

$$D^{cc}(f \circ g(AP_q, c)) \geq \frac{\log_2(q/n) - 200}{4} \cdot D^{dt}(f)$$

(in  $g(AP_q, c)$  each party receives  $2 \log_2 q$  bits).

We also give an example of a natural-looking gadget for which Corollary 3 implies a simulation theorem. Our gadget is the following one: Alice gets  $a \in \mathbb{F}_{q^2}$  and Bob gets  $b \in \mathbb{F}_{q^2}$ . Here  $q$  is a power of an odd prime. Their goal is to output 1, if  $a - b$  is a square in  $\mathbb{F}_{q^2}$  (by that we mean that there is  $c \in \mathbb{F}_{q^2}$  such that  $a - b = c^2$ ), and 0 otherwise. Let us denote this gadget by  $\text{SQR}^q$ .

Since  $\mathbb{F}_{q^2}$  is a linear space over  $\mathbb{F}_q$ , we can naturally identify inputs to  $\text{SQR}^q$  with  $\mathbb{F}_q^2$ , i.e.  $\text{SQR}^q$  can be viewed as a function of the form  $\text{SQR}^q : \mathbb{F}_q^2 \times \mathbb{F}_q^2 \rightarrow \{0, 1\}$ .

► **Proposition 4.** *For all large enough  $q$  the following holds. If  $q$  is a power of an odd prime, then there exists a balanced covering  $c$  of  $AP_q$  such that  $g(AP_q, c)$  is a sub-function of  $\text{SQR}^q$ , i.e. whenever  $g(AP_q, c)(a, b)$  is defined, we have  $g(AP_q, c)(a, b) = \text{SQR}^q(a, b)$ . A truth table of  $c$  can be computed in time  $q^{O(1)}$ .*

This Proposition implies a simulation theorem for  $\text{SQR}^q$ , with the same parameters as in Corollary 3 and with polynomial-time listable underlying hitting distributions.

Next we observe that any spectral expander “similar” to  $AP_q$  automatically satisfies restrictions of Theorem 2.

► **Proposition 5.** *Assume that  $G = (V, E)$  is a  $(m, d, \gamma)$ -spectral expander and*

$$2d + 4 > d^2 \left( 2\gamma^2 + \frac{4(1 - \gamma^2)}{m} \right).$$

*Then for any two distinct vertices  $u, v \in V$  it holds that  $|\Gamma(u) \cap \Gamma(v)| \leq 1$ .*

In particular, all  $(m^2, m, 1/\sqrt{m})$ -spectral expanders satisfy these restrictions. However, Proposition 5 is by no means a necessary condition. For example, Theorem 2 can be also applied Lubotzky-Phillips-Sarnak construction of Ramanujan graphs ([8]). More specifically, if  $p, q$  are unequal primes,  $p, q \equiv 1 \pmod{4}$  and  $p$  is a quadratic residue modulo  $q$ , the paper [8] constructs an explicit graph  $X^{p,q}$  which, in particular, is a  $(q(q^2 - 1)/2, p + 1, 2\sqrt{p}/(p + 1))$ -spectral expander and in which the shortest cycle is of length at least  $2 \log_p q$ . It can also be easily shown that provided  $p < q^2$  there are no self-loops in  $X^{p,q}$ . Thus if  $p < \sqrt{q}$ , then any two distinct vertices of  $X^{p,q}$  have at most one common neighbor, while inequality from Proposition 5 is false for  $X^{p,q}$ .

We then obtain some results related to the following question: what is the best possible trade-off between the arity of outer functions and the size of gadget in deterministic simulation theorems? Once again, consider  $\text{SQR}^q$ . Note that in  $\text{SQR}^q$  each party receives  $k = 2 \log_2 q$  bits. Corollary 3 lower bounds  $D^{cc}(f \circ \text{SQR}^q)$  whenever arity of  $f$  is at most  $2^{k/2 - O(1)}$ . In this regime the lower bound is of the form  $\Omega(D^{dt}(f))$  (compare it to the  $O(k \cdot D^{dt}(f))$  upper bound). In turn, if the arity of  $f$  is at most  $2^{(1/2 - \Omega(1))k}$ , the lower bound becomes tight, namely  $\Omega(k \cdot D^{dt}(f))$ . Thus  $\text{SQR}^q$  achieves the same trade-off between the arity of  $f$  and the size of a gadget as  $k$ -bit Inner Product (while underlying hitting distributions for  $\text{SQR}^q$ , unlike Inner Product, are polynomial-time listable).

Ramanujan graphs yield gadgets with much worse trade-off. Namely, if  $p$  is of order  $\sqrt{q}$  and  $c$  is a balanced coloring of  $X^{p,q}$ , then  $g(X^{p,q}, c)$  is a gadget on  $k \approx 3 \log_2 q$  bits which admits a simulation theorem for all outer functions of arity roughly  $2^{\log_2 p} = 2^{k/6}$ .

This raises the following question: for a given  $k$  what is the maximal  $h$  such that there is a gadget on  $k$  bits having two  $(\frac{1}{10}, h)$ -hitting distributions, the one over 0-monochromatic rectangles and the other over 1-monochromatic rectangles? Above discussion shows that  $h$  can be about  $k/2$ . In the following Proposition we observe that it is impossible to obtain any constant bigger than  $1/2$  before  $k$ .

► **Proposition 6.** *For every  $g : \{0, 1\}^k \times \{0, 1\}^k \rightarrow \{0, 1\}$  and for every integer  $h \geq 1$  there exists  $b \in \{0, 1\}$  such that the following holds. For every probability distribution  $\mu$  over  $b$ -monochromatic rectangles of  $g$  there are  $X, Y \subset \{0, 1\}^k$  of size at least  $2^{k-h}$  such that*

$$\Pr_{R \sim \mu} [R \cap X \times Y \neq \emptyset] \leq 2^{k-2h+1}.$$

In addition we show the following simple proposition, studying the minimal possible support size of hitting distributions.

► **Proposition 7.** *For every  $g : \{0, 1\}^k \times \{0, 1\}^k \rightarrow \{0, 1\}$  the following holds*

- *if there is  $(\frac{1}{20}, h)$ -hitting distribution over  $b$ -monochromatic rectangles of  $g$  for some  $b \in \{0, 1\}$ , then there is  $(\frac{1}{10}, h)$ -hitting distribution over  $b$ -monochromatic rectangles of  $g$  which support is of size  $2^{O(k)}$ .*
- *Assume that for some  $\delta < 1$  and  $h \in \mathbb{N}$  there are two  $(\delta, h)$ -hitting distributions  $\mu_0, \mu_1$ , where  $\mu_b$  is over  $b$ -monochromatic rectangles of  $g$ . Then the support of  $\mu_b$  is of size at least  $2^h$ , for every  $b \in \{0, 1\}$ .*

So it is impossible to improve a trade-off between the size of outer functions and the size of gadgets simply by improving hitting distributions. However, until now we only spoke about improving gadgets. What about outer functions? What causes a restriction on the arity of  $f$  in Theorem 1? It can be verified that the only place in which arity of  $f$  appears in the proof is so-called Thickness Lemma. Let us state this Lemma.

Assume that  $A$  is a finite set and  $X$  is a subset of  $A^n$ . Here  $n$  corresponds to the arity of  $f$ . Let  $X_{[n]/\{i\}}$  denote the projection of  $X$  onto all the coordinates except the  $i$ -th one. Define the following auxiliary bipartite graph  $G_i(X)$ . Left side vertices of  $G_i(X)$  are taken from  $A$ , right side vertices of  $G_i(X)$  are taken from  $X_{[n]/\{i\}}$ . We connect  $a \in A$  with  $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in X_{[n]/\{i\}}$  if and only if

$$(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \in X.$$

Clearly, there are  $|X|$  edges in  $G_i(X)$ .

Let  $MinDeg_i(X)$  denote the minimal possible degree of a right side vertex of  $G_i(X)$ . Similarly, let  $AvgDeg_i(X)$  denote the average degree of a right side vertex of  $G_i(X)$ . There are  $|X|$  edges and  $|X_{[n]/\{i\}}|$  right side vertices, hence it is naturally to define  $AvgDeg_i(X)$  as

$$AvgDeg_i(X) = \frac{|X|}{|X_{[n]/\{i\}}|}.$$

Thickness Lemma relates this two measures. Namely, it states that if for every  $i$  average degree of  $G_i(X)$  is big, then there is a large subset  $X' \subset X$  such that for every  $i$  minimal degree of  $G_i(X')$  is big. The precise bounds can be found in the following

► **Lemma 8** ([9]). *Consider any  $\delta \in (0, 1)$ . Assume that for every  $i \in \{1, 2, \dots, n\}$  we have that  $AvgDeg_i(X) \geq d$ . Then there is  $X' \subset X$  of size at least  $(1 - \delta)|X|$  such that for every  $i \in [n]$  it holds that  $MinDeg_i(X') \geq \frac{\delta d}{n}$ .*

One possible way to improve a trade-off between the arity of  $f$  and the size of gadget is to improve Thickness Lemma. For example, if we could replace  $\frac{\delta d}{n}$  with  $\frac{\delta d}{\sqrt{n}}$  in Lemma 8, this would mean that  $k$ -bit Inner Product and  $k$ -bit SQR-gadget admit simulation theorems for all outer functions of arity roughly  $2^k$  (rather than  $2^{k/2}$ ).

However, such an improvement is impossible and the bounds given in Lemma 8 are near-optimal. Note that Thickness Lemma says nothing about whether there even exists a non-empty subset  $X' \subset X$  such that for all  $i \in [n]$  it holds that  $MinDeg_i(X')$  is larger, say, by a constant than  $\frac{d}{n}$ . And indeed, we show that for some  $X$  there is no such  $X'$  at all. More precisely, we show the following

► **Theorem 9.** *For every  $\varepsilon > 0$  and for all integer  $n \geq 2, s \geq 1$  there exists  $m \in \mathbb{N}$  and a non-empty set  $X \subset \{0, 1, \dots, m-1\}^n$  such that*

- *for all  $i \in [n]$  it holds that  $AvgDeg_i(X) \geq s(n - \varepsilon)$ ;*
- *there is no non-empty  $Y \subset X$  such that for all  $i \in [n]$  it holds that  $MinDeg_i(Y) \geq s + 1$ .*

## 1.2 Organization of the paper

The rest of the paper is organized as follows.

In Section 2 we give Preliminaries. In Section 3 we prove Theorem 2 and derive Corollary 3. In Section 4 we prove Proposition 4. In Section 5 we prove Theorem 9. In Section 6 we prove Proposition 5. In section 7 we prove proposition 6. The proof of Proposition 7 is omitted due to space constraints.

## 2 Preliminaries

### 2.1 Sets notations

Let  $[n]$  be the set  $\{1, 2, \dots, n\}$ .

Assume that  $A$  is a finite set,  $X$  is a subset of  $A^n$  and  $S = \{i_1, \dots, i_k\}$ , where  $i_1 < i_2 < \dots < i_k$ , is a subset of  $[n]$ . Let  $X_S$  denote the following set:

$$X_S = \{(x_{i_1}, \dots, x_{i_k}) : (x_1, \dots, x_n) \in X\} \subset A^{|S|}.$$

Given  $X \subset A^n$  and  $i \in [n]$ , consider the following bipartite graph  $G_i(X) = (A, X_{[n] \setminus \{i\}}, E)$ , where

$$E = \{(x_i, (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)) : (x_1, \dots, x_n) \in X\}.$$

Vertices of  $G_i(X)$  which are from  $A$  will be called left vertices. Similarly, vertices of  $G_i(X)$  which are from  $X_{[n] \setminus \{i\}}$  will be called right vertices.

Define  $MinDeg_i(X)$  as minimal  $d$  such that there is a right vertex of  $G_i(X)$  with degree  $d$ . Define  $AvgDeg_i(X) = |X|/|X_{[n] \setminus \{i\}}|$ .

### 2.2 Communication and query complexity

For introduction in both query and communication complexities see, e.g., [7]. We will use the following notation.

For a Boolean function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  let  $D^{dt}(f)$  denote  $f$ 's deterministic query complexity, i.e. minimal  $d$  such that there is a deterministic decision tree of depth  $d$  computing  $f$ . For a (possibly partial) Boolean function  $g : A \times B \rightarrow \{0, 1\}$ , where  $A, B$  are some finite sets, let  $D^{cc}(g)$  denote  $g$ 's deterministic communication complexity, i.e. minimal  $d$  such that there is a deterministic communication protocol of depth  $d$ , computing  $g$ . Let us stress that in the case when  $g$  is partial by “deterministic communication protocol computes  $g$ ” we mean only that a protocol outputs 0 on  $(a, b)$  whenever  $g(a, b) = 0$  and outputs 1 on  $(a, b)$  whenever  $g(a, b) = 1$ ; on inputs on which  $g$  is not defined the protocol may output anything.

If  $f, g$  are as above, let  $f \circ g$  denote the following (possibly partial) function:

$$\begin{aligned} f \circ g : A^n \times B^n &\rightarrow \{0, 1\}, \\ (f \circ g)((a_1, \dots, a_n), (b_1, \dots, b_n)) &= f(g(a_1, b_1), \dots, g(a_n, b_n)). \end{aligned}$$

We can also measure  $D^{cc}(f \circ g)$ , deterministic communication complexity of  $f \circ g$ , assuming that Alice's input is  $(a_1, \dots, a_n) \in A^n$  and Bob's input is  $(b_1, \dots, b_n) \in B^n$ .

### 2.3 Hitting distributions

Fix a (possibly partial) Boolean function  $g : A \times B \rightarrow \{0, 1\}$ . A set  $R \subset A \times B$  is called *rectangle* if there are  $U \subset A, V \subset B$  such that  $R = U \times V$ . If  $b \in \{0, 1\}$ , then we say that rectangle  $R$  is  $b$ -monochromatic for  $g$  if  $g(a, b) = b$  whenever  $(a, b) \in R$ . We stress that if  $g$  is partial, then in the definition of  $b$ -monochromatic rectangle we require that  $g$  is everywhere defined on  $R$ .

Let  $\delta$  be positive real and  $h$  be positive integer. A probability distribution  $\mu$  over rectangles  $R \subset A \times B$  is called  $(\delta, h)$ -hitting if for all  $X \subset A, Y \subset B$  such that  $|X| \geq 2^{-h}|A|, |Y| \geq 2^{-h}|B|$  it holds that

$$\Pr_{R \sim \mu} [R \cap X \times Y \neq \emptyset] \geq 1 - \delta.$$

In this paper we are focused only on those  $\mu$  such that there exists  $b \in \{0, 1\}$  for which all rectangles from the support of  $\mu$  are  $b$ -monochromatic for  $g$ . In this case we simply say that  $\mu$  is over  $b$ -monochromatic rectangles of  $g$ .

Let  $g_t : \{0, 1\}^{k_t} \times \{0, 1\}^{k_t} \rightarrow \{0, 1\}$  be family of gadgets and  $\mu_t$  be family of probability distributions, where  $\mu_t$  is over rectangles of  $g_t$ . We call  $\mu_t$  polynomial-time listable if the following holds:

- the size of the support of  $\mu_t$  is  $2^{O(k_t)}$ ;
- all the probabilities of  $\mu_t$  are rational;
- there is a deterministic Turing machine which, having  $k_t$  on input, in time  $2^{O(k_t)}$  computes  $g_t$ 's matrix and lists all the rectangles from the support of  $\mu_t$ , together with probabilities  $\mu_t$  assigns to them.

## 2.4 SQR-gadget

Consider a finite field of size  $q$ , denoted below by  $\mathbb{F}_q$ . We call  $a \in \mathbb{F}_q$  a *square* if there is  $b \in \mathbb{F}_q$  such that  $a = b^2$  in  $\mathbb{F}_q$ . Let  $\text{SQR}^q$  denote the following Boolean function:

$$\text{SQR}^q : \mathbb{F}_{q^2} \times \mathbb{F}_{q^2} \rightarrow \{0, 1\}, \quad \text{SQR}^q(a, b) = \begin{cases} 1 & \text{if } a - b \text{ is a square in } \mathbb{F}_{q^2}, \\ 0 & \text{if } a - b \text{ is not a square in } \mathbb{F}_{q^2}. \end{cases}$$

## 2.5 Expanders

We consider undirected graphs which may have parallel edges and self-loops. We assume that a self-loop at vertex  $v$  contributes 1 to degree of  $v$ . A graph is called  $d$ -regular if each its vertex has degree  $d$ .

A coloring of a graph  $G = (V, E)$  is a function  $c : V \rightarrow \{0, 1\}$ . It is called balanced if  $|V|/3 \leq |c^{-1}(1)| \leq 2|V|/3$ . For any  $A \subset V$  let  $\Gamma(A)$  denote the set of all  $v \in V$  such that there is  $u \in A$  connected with  $v$  by an edge of  $G$ . If  $v \in V$ , define  $\Gamma(v) = \Gamma(\{v\})$ .

Fix graph  $G = (V, E)$  and a coloring  $c : V \rightarrow \{0, 1\}$ . Assume that for any two distinct  $u, v \in V$  it holds that  $|\Gamma(u) \cap \Gamma(v)| \leq 1$ . Then the following partial function is well defined:

$$g(G, c) : V \times V \rightarrow \{0, 1\},$$

$$g(G, c)(u, v) = \begin{cases} 1 & u \neq v \text{ and there is } w \in \Gamma(u) \cap \Gamma(v) \text{ s.t. } c(w) = 1, \\ 0 & u \neq v \text{ and there is } w \in \Gamma(u) \cap \Gamma(v) \text{ s.t. } c(w) = 0, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Let  $M_G$  be an adjacency matrix of a  $d$ -regular graph  $G = (V, E)$  with  $|V| = m$ . Note that  $d$  is an eigenvalue of  $M_G$ . A graph  $G$  is called  $(m, d, \gamma)$ -spectral expander if  $M_G$  satisfies the following conditions:

- multiplicity of an eigenvalue  $d$  is 1;
- absolute value of any other eigenvalue of  $M_G$  is at most  $\gamma d$ .

► **Proposition 10** ([11], Theorem 4.6). *Assume that a graph  $G = (V, E)$  is  $(m, d, \gamma)$ -spectral expander. Then for any  $A \subset V$ :*

$$\frac{|\Gamma(A)|}{|A|} \geq \frac{1}{\gamma^2 + (1 - \gamma^2) \frac{|A|}{m}}.$$



Assume the  $q$  is a power of prime. Let  $AP_q$  denote the following graph. Vertices of  $AP_q$  are pairs of elements of  $\mathbb{F}_q$  so that the number of vertices is  $q^2$ . We connect  $(x, y)$  with  $(a, b)$  by an edge if and only if  $ax = b + y$  in  $\mathbb{F}_q$ . It is easy to see that  $AP_q$  is  $q$ -regular.

► **Proposition 11** ([10], Lemma 5.1).  $AP_q$  is  $(q^2, q, 1/\sqrt{q})$ -spectral expander.

## 2.6 $k$ -wise independent hash functions

We will need the following

► **Proposition 12** ([11], Corollary 3.34). For every  $n, k \in \mathbb{N}$  there exists a polynomial-time computable function  $\psi : \{0, 1\}^{kn} \times \{0, 1\}^n \rightarrow \{0, 1\}$  such for all distinct  $x_1, \dots, x_k \in \{0, 1\}^n$  and for all  $b_1, \dots, b_k \in \{0, 1\}$  the following holds:

$$\Pr[\psi(s, x_1) = b_1, \dots, \psi(s, x_k) = b_k] = 2^{-k},$$

where the probability is over uniformly random  $s \in \{0, 1\}^{kn}$ .

## 2.7 Some useful facts

We will use the following inequality involving binomial coefficients:

► **Lemma 13.** For every  $k, m$  the following holds: if  $k \leq m/2$ , then  $\binom{m-k}{k} / \binom{m}{k} \geq 1 - \frac{k^2}{m-k}$ .

**Proof.** Observe that:

$$\begin{aligned} \frac{\binom{m-k}{k}}{\binom{m}{k}} &= \frac{m-k}{m} \cdot \frac{m-k-1}{m-1} \cdot \dots \cdot \frac{m-2k+1}{m-k+1} \geq \left(\frac{m-2k}{m-k}\right)^k \\ &= \left(1 - \frac{k}{m-k}\right)^k \geq 1 - \frac{k^2}{m-k}. \end{aligned}$$

The first inequality here is due to the fact that for all positive  $i$  we have:

$$\begin{aligned} \frac{k}{m-k+i} \leq \frac{k}{m-k} &\implies 1 - \frac{k}{m-k+i} \geq 1 - \frac{k}{m-k} \\ &\implies \frac{m-2k+i}{m-k+i} \geq \frac{m-2k}{m-k}. \end{aligned}$$

The second inequality here is Bernoulli's inequality. It is legal to apply this inequality because  $k \leq m/2$  and hence  $\frac{k}{m-k} \leq 1$ . ◀

Note that  $\mathbb{F}_{q^2}$  contains a subfield of size  $q$ . Namely,  $\mathbb{F}_q = \{x \in \mathbb{F}_{q^2} : x^q = x\}$ .

► **Lemma 14.** Assume that  $q$  is a power of an odd prime. Let  $\alpha$  be a primitive root of  $\mathbb{F}_{q^2}$ . Then the following holds:

- $0, \alpha^2, \alpha^4, \dots, \alpha^{q^2-1}$  are the only squares in  $\mathbb{F}_{q^2}$ ;
- all the elements of  $\mathbb{F}_q$  are squares in  $\mathbb{F}_{q^2}$ .

**Proof.** Let us prove the first statement of the lemma. Assume that  $j \in \{1, 2, \dots, q^2 - 1\}$  is an odd integer. We will show that  $\alpha^j$  is not a square. Indeed, assume for contradiction there is a non-zero  $y \in \mathbb{F}_{q^2}$  such that  $\alpha^j = y^2$ . Therefore for some integer  $i$  we have that  $\alpha^{j-2i} = 1$ . Since  $\alpha$  is the primitive root of  $\mathbb{F}_{q^2}$ , this means that  $j - 2i$  is divisible by  $q^2 - 1$ . But  $j - 2i$  is odd and  $q^2 - 1$  is even.

To show the second statement of the lemma assume that  $x = \alpha^k$  is a non-zero root of  $x^q = x$ . Then we have that  $\alpha^{k(q-1)} = 1$ . Due to the same argument as above  $k(q-1)$  is divisible by  $q^2 - 1$ . This implies that  $k$  is divisible by  $q+1$ . Hence  $k$  is even and  $x = \alpha^k$  is a square. ◀

### 3 Transforming Expanders into Gadgets

In this section we prove Theorem 2 and derive Corollary 3.

**Proof of Theorem 2.** Fix  $b \in \{0, 1\}$  and set  $h = \lfloor 2 \log_2(1/\gamma) \rfloor - 100$ . Let us define a  $(\frac{1}{10}, h)$ -hitting distribution  $\mu_b$  over  $b$ -monochromatic rectangles of  $g(G, c)$ . Take  $v \in c^{-1}(b)$  uniformly at random. Split  $\Gamma(v)$  into two disjoint subsets  $A, B$  randomly according to 10-wise independent hash function  $\psi : \{0, 1\}^{10 \cdot \lceil \log_2 m \rceil} \times \{0, 1\}^{\lceil \log_2 m \rceil} \rightarrow \{0, 1\}$  from Proposition 12. Namely, take  $s \in \{0, 1\}^{10 \lceil \log_2 m \rceil}$  uniformly at random. An element  $u \in \Gamma(v)$  goes into  $A$  if  $\psi(s, u) = 0$  and into  $B$  if  $\psi(s, u) = 1$ . By definition  $A \times B$  is a  $b$ -monochromatic rectangle of  $g(G, c)$ . Indeed, any two distinct vertices from  $\Gamma(v)$  have a common neighbor colored in  $b$ . It remains to show that for all  $S, T \subset V$  of size at least  $2^{-h}m$  with probability at least 0.9 we have that  $A \times B \cap S \times T \neq \emptyset$ . It is enough to show that  $\Pr[A \cap S \neq \emptyset] \geq 0.96$  and  $\Pr[B \cap T \neq \emptyset] \geq 0.96$ . Let us show that the first inequality holds, the proof of the second inequality is exactly the same. Actually we will show that  $\Pr[|\Gamma(v) \cap S| \geq 10] \geq 0.97$ . This is enough for our purposes: conditioned on  $|\Gamma(v) \cap S| \geq 10$  the probability that  $A$  is disjoint with  $S$  is at most  $2^{-10}$  (due to proposition 12 this is the probability that  $\psi(s, \cdot)$  sends 10 fixed points of  $\Gamma(v)$  into  $B$ ). Therefore  $\Pr[A \cap S] \geq (1 - 2^{-10}) \Pr[|\Gamma(v) \cap S| \geq 10] \geq 0.999 \cdot 0.97 > 0.96$ .

The size of  $S$  is at least  $2^{100}\gamma^2 m$ . Partition  $S$  into 10 disjoint subsets  $S_1, \dots, S_{10}$ , each of size at least  $2000\lceil \gamma^2 m \rceil$ . Since  $m \geq 1/\gamma^2$ , we also have  $|S_1|, \dots, |S_{10}| \geq 1000\gamma^2 m$ . If  $|\Gamma(v) \cap S| < 10$ , then  $\Gamma(v)$  is disjoint with  $S_i$  for some  $i \in [10]$ . Hence

$$\Pr[|\Gamma(v) \cap S| < 10] \leq \sum_{i=1}^{10} \Pr[\Gamma(v) \cap S_i = \emptyset].$$

If we show for all  $i \in [10]$  that  $\Pr[\Gamma(v) \cap S_i = \emptyset] \leq 0.003$ , we are done. Observe that  $\Gamma(v)$  is disjoint with  $S_i$  if and only if  $v \notin \Gamma(S_i)$ . This implies that

$$\Pr[\Gamma(v) \cap S_i = \emptyset] = \frac{|c^{-1}(b) \setminus \Gamma(S_i)|}{|c^{-1}(b)|} \leq \frac{m - |\Gamma(S_i)|}{\frac{m}{3}}. \quad (2)$$

In the last inequality we use the fact that  $c$  is balanced. By Proposition 10 we get

$$|\Gamma(S_i)| \geq \frac{|S_i|}{\gamma^2 + \frac{|S_i|}{m}} \geq \frac{|S_i|}{\frac{|S_i|}{1000 \cdot m} + \frac{|S_i|}{m}} \geq \frac{1000 \cdot m}{1001} > 0.999m.$$

Here in the second inequality we use the fact that  $|S_i| \geq 1000\gamma^2 m$ . Due to (2) this means that  $\Pr[\Gamma(v) \cap S_i = \emptyset] \leq 0.003$  and thus the proof that  $\mu_b$  is  $(\frac{1}{10}, h)$ -hitting is finished.

Let us now show that  $\mu_b$  can be “written down” in time  $m^{O(1)}$  from  $G$  and  $c$ . First of all, note that  $g(G, c)$  is a gadget on  $k = \lceil \log_2 m \rceil$  bits. To specify a rectangle from a support of  $\mu_b$  we need to specify a vertex of  $G$  and a “seed”  $s$  of length  $10k$ . This shows that the support of  $\mu_b$  is of size  $m^{O(1)} = 2^{O(k)}$ . This observation also allows us to list all the rectangles from the support of  $\mu_b$  in time  $2^{O(k)}$  – just go through all vertices from  $c^{-1}(b)$  and all seeds. Further, the  $\mu_b$ -probability of  $A \times B$  can be computed as follows:

$$\mu_b(A \times B) = \frac{|\{v \in V : \Gamma(v) = A \cup B\}| \cdot |\{s \in \{0, 1\}^{10k} : \phi(s, \cdot) \text{ splits } A \cup B \text{ into } A \text{ and } B\}|}{|c^{-1}(b)| \cdot 2^{10k}}.$$

This probability is rational and can be computed in time  $2^{O(k)}$ , again by exhaustive search over all vertices and seeds.  $\blacktriangleleft$

Now let us derive Corollary 3. Indeed,  $AP_q$  is  $(q^2, q, 1/\sqrt{q})$ -spectral expander by Proposition 11. Thus theorem 2, applied to  $AP_q$ , states that for any balanced coloring  $c$  of  $AP_q$  and for any  $b \in \{0, 1\}$  there exists  $(\frac{1}{10}, \lfloor \log_2(q) \rfloor - 100)$ -hitting distribution over  $b$ -monochromatic rectangles of  $g(AP_q, c)$ . Apply Theorem 1 to these hitting distributions with  $\varepsilon = 1 - \log_2(n)/(\lfloor \log_2(q) \rfloor - 100)$ .

We only need to check that in  $AP_q$  for any two distinct vertices  $u, v$  it holds that  $|\Gamma(u) \cap \Gamma(v)| \leq 1$ . Assume that  $(x, y)$  and  $(u, v)$  are distinct vertices of  $AP_q$ . Take any  $(a, b) \in \Gamma((x, y)) \cap \Gamma((u, v))$ . Then

$$\begin{pmatrix} x & -1 \\ u & -1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y \\ v \end{pmatrix}. \quad (3)$$

If  $x \neq u$ , then  $\det \begin{pmatrix} x & -1 \\ u & -1 \end{pmatrix} \neq 0$  and hence system (3) has exactly one solution. If  $x = u$ , then  $y \neq v$  and system (3) has no solution. Therefore  $|\Gamma((x, y)) \cap \Gamma((u, v))| \leq 1$ .

## 4 SQR<sup>q</sup> Gadget

In this section we prove Proposition 4.

Fix  $w \in \mathbb{F}_{q^2}$  such that  $\{1, w\}$  is a basis of  $\mathbb{F}_{q^2}$  over  $\mathbb{F}_q$ . Consider the following coloring of  $AP_q$ : set  $c((a, b)) = 1$  if and only if  $1 + wa$  is a square in  $\mathbb{F}_{q^2}$ ; clearly a truth table of such  $c$  can be computed in time  $q^{O(1)}$ . Note that  $g(AP_q, c)((x, y), (u, v))$  is defined if and only if  $(x, y), (u, v)$  are distinct and there is  $(a, b) \in \Gamma((x, y)) \cap \Gamma((u, v))$ . Let us show that for any such  $(x, y), (u, v)$  it holds that

$$g(AP_q, c)((x, y), (u, v)) = c((a, b)) = \text{SQR}^q(x + yw, u + vw). \quad (4)$$

Indeed, we have that  $ax = b + y, au = b + v$ . This means that  $y - v = a(x - u)$ . Moreover, due to distinctness of  $(x, y), (u, v)$  we have that  $x \neq u$ . Further,

$$x + yw - (u + vw) = (x - u) + w(y - v) = (x - u)(1 + wa).$$

Note that  $x - u$  is a non-zero element of  $\mathbb{F}_q$ . By the second item of Lemma 14 this implies that  $x + yw - (u + vw)$  is a square if and only if  $1 + wa$  is a square. Hence (4) is true for all  $(x, y), (u, v)$  from the domain of  $g(AP_q, c)$ .

It remains to show that  $c$  is balanced. Take  $(a, b, \lambda) \in \mathbb{F}_q \times \mathbb{F}_q \times (\mathbb{F}_q \setminus \{0\})$  uniformly at random. Note that  $c((a, b)) = 1$  if and only if  $1 + wa$  is a square. Thus  $|c^{-1}(1)| = q^2 \Pr[1 + wa \text{ is a square}]$ . Due to the second item of Lemma 14 we have that  $1 + wa$  is a square if and only if  $\lambda(1 + wa)$  is a square. Note that  $\lambda(1 + wa) = \lambda + \lambda aw$  is distributed uniformly in  $\{i + wj : i, j \in \mathbb{F}_q, i \neq 0\}$  (this is because for any  $\lambda_0$  the distribution of  $\lambda a$  given  $\lambda = \lambda_0$  is uniform in  $\mathbb{F}_q$ ). Due to the first item of Lemma 14 for all large enough  $q$  there are at least  $0.4q^2$  squares and at least  $0.4q^2$  non-squares in  $\{i + wj : i, j \in \mathbb{F}_q, i \neq 0\}$ . This means that  $1/3 \leq \Pr[\lambda(1 + wa) \text{ is a square}] \leq 2/3$  for all large enough  $q$ . Hence  $q^2/3 \leq |c^{-1}(1)| \leq 2q^2/3$  and  $c$  is balanced.

## 5 Unimprovability of Thickness Lemma

Consider any set  $X \subset \{0, 1, \dots, m-1\}^n$  and take any  $i \in [n]$ . Let us say that  $x \in X$  is  $i$ -unique in  $X$  if there is no other  $x' \in X$  such that

$$x_1 = x'_1, \dots, x_{i-1} = x'_{i-1}, x_{i+1} = x'_{i+1}, \dots, x_n = x'_n.$$

Call a set  $X \subset \{0, 1, \dots, m-1\}^n$  *reducible* if for all non-empty  $Y \subset X$  there is  $i \in [n]$  such that  $\text{MinDeg}_i(Y) = 1$ . Note that  $X$  is reducible if and only if for all non-empty  $Y \subset X$  there is  $y \in Y$  which is  $i$ -unique in  $Y$  for some  $i \in [n]$ .

► **Lemma 15.** *For every  $\varepsilon > 0$  and for every  $n \geq 2$  there exists  $m > 0$  and a reducible set  $X \subset \{0, 1, \dots, m-1\}^n$  such that for all  $i \in [n]$  it holds that  $\text{AvgDeg}_i(X) \geq n - \varepsilon$ .*

**Proof.** Take any  $m > 0$ . Consider the following sequence of sets  $X_2, X_3, \dots$ , where  $X_n$  is a subset of  $\{0, 1, \dots, m-1\}^n$ :

$$X_2 = \{(j, j) : j \in \{0, 1, \dots, m-1\}\} \cup \{(j, j+1) : j \in \{0, 1, \dots, m-2\}\},$$

$$X_{\ell+1} = \{(x, j) : x \in X_\ell, j \in \{0, 1, \dots, m-1\}\} \\ \cup \{(y, 0) : y \in \{0, 1, \dots, m-1\}^\ell / X_\ell\}.$$

We have the following relation between the size of  $X_{\ell+1}$  and the size of  $X_\ell$ :

$$|X_{\ell+1}| = m \cdot |X_\ell| + m^\ell - |X_\ell| = m \cdot (|X_\ell| - 1) + m^\ell.$$

Let us show by induction on  $n$  that  $|X_n| \geq nm^{n-1} - n(1 + m + \dots + m^{n-2})$ . Indeed, for  $n = 2$  this inequality is true:  $|X_2| = 2m - 1 > 2m - 2$ . Now, assume that for  $n = \ell$  this inequality is proved, i.e.  $|X_\ell| \geq \ell m^\ell - \ell(1 + m + \dots + m^{\ell-2})$ . Then

$$|X_{\ell+1}| = m \cdot (|X_\ell| - 1) + m^\ell \\ \geq m \cdot (\ell m^{\ell-1} - \ell(1 + m + \dots + m^{\ell-2}) - 1) + m^\ell \\ \geq (\ell + 1) \cdot m^\ell - (\ell + 1) \cdot (1 + m + \dots + m^{\ell-1}).$$

This means that for every  $n$  and  $i \in [n]$  it holds that

$$\text{AvgDeg}_i(X_n) = \frac{|X_n|}{|(X_n)_{[n]/\{i\}}|} \geq \frac{nm^{n-1} - n(1 + m + \dots + m^{n-2})}{m^{n-1}},$$

and the latter tends to  $n$  as  $m \rightarrow \infty$ . Thus to show the lemma it is sufficient to show that  $X_n$  is reducible. Once again, we will show it by induction on  $n$ .

Consider  $n = 2$  and take any non-empty  $Y \subset X_2$ . Let  $y \in Y$  be the smallest element of  $Y$  in lexicographical order. If  $y = (j, j)$ , then  $y$  is 1-unique in  $Y$  and hence  $\text{MinDeg}_1(Y) = 1$ . If  $y = (j, j+1)$ , then  $y$  is 2-unique in  $Y$  and hence  $\text{MinDeg}_2(Y) = 1$ .

Further, assume that for  $n = \ell$  the statement is proved, i.e.  $X_\ell$  is reducible. Consider any non-empty  $Y \subset X_{\ell+1}$ . Assume that  $Y$  intersects  $\{(y, 0) : y \in \{0, 1, \dots, m-1\}^\ell / X_\ell\}$  and hence for some  $y \notin X_\ell$  it holds that  $(y, 0) \in Y$ . Then  $\text{MinDeg}_{\ell+1}(Y) = 1$ . Indeed, in this case  $(y, 0)$  is  $(\ell + 1)$ -unique in  $Y$ , because if  $(y, j) \in Y \subset X_{\ell+1}$  for some  $j > 0$ , then  $y \in X_\ell$ , contradiction.

Now assume that  $Y$  is a subset of  $\{(x, j) : x \in X_\ell, j \in \{0, 1, \dots, m-1\}\}$ . Then for some  $j \in \{0, 1, \dots, m-1\}$  a set  $Y' = \{x \in X_\ell : (x, j) \in Y\}$  is non-empty. Since by induction hypothesis  $X_\ell$  is reducible, there is  $y \in Y'$  which is  $i$ -unique in  $Y'$  for some  $i \in [\ell]$ . Let us show that  $(y, j)$  is  $i$ -unique in  $Y$  (this would mean that  $\text{MinDeg}_i(Y) = 1$ ). Indeed, assume that there is  $(y', j') \in Y$  which coincides with  $(y, j)$  on all the coordinates except the  $i^{\text{th}}$  one. Then  $j = j'$  and  $y' \in Y'$ . Due to  $i$ -uniqueness of  $y \in Y'$  we also have that  $y = y'$ . ◀

► **Definition 16.** Let  $s, m, n$  be positive integers and assume that  $X$  is a subset of  $\{0, 1, \dots, m-1\}^n$ . Let  $\text{In}(X, s) \subset \{0, 1, \dots, sm-1\}^n$  denote the following set:

$$\text{In}(X, s) = \{(sx_1 + r_1, sx_2 + r_2, \dots, sx_n + r_n) : \\ (x_1, \dots, x_n) \in X, r_1, \dots, r_n \in \{0, 1, \dots, s-1\}\}.$$

Observe that for every  $(y_1, \dots, y_n) \in \text{In}(X, s)$  there is exactly one  $(x_1, \dots, x_n) \in X$  such that for some  $r_1, \dots, r_n \in \{0, 1, \dots, s-1\}$  it holds that

$$y_1 = sx_1 + r_1, \dots, y_n = sx_n + r_n.$$

► **Lemma 17.** *For every  $i \in \{1, 2, \dots, n\}$  it holds that  $\text{AvgDeg}_i(\text{In}(X, s)) = s \cdot \text{AvgDeg}_i(X)$ .*

**Proof.** Lemma follows from the following two equalities:

$$|\text{In}(X, s)| = s^n \cdot |X|, \quad |\text{In}(X, s)_{[n]/\{i\}}| = s^{n-1} \cdot |X_{[n]/\{i\}}|. \quad \blacktriangleleft$$

► **Lemma 18.** *Assume that  $X \subset \{0, 1, \dots, m-1\}^n$  is reducible. Then for all non-empty  $Y \subset \text{In}(X, s)$  there is  $i \in [n]$  such that  $\text{MinDeg}_i(Y) \leq s$ .*

**Proof.** Let  $Y'$  be the set of all  $(x_1, \dots, x_n) \in X$  for which there are

$$r_1, \dots, r_n \in \{0, 1, \dots, s-1\}$$

such that  $(sx_1 + r_1, \dots, sx_n + r_n) \in Y$ . Clearly  $Y'$  is non-empty. Let  $x' = (x'_1, \dots, x'_n) \in Y'$  be a string which is  $i$ -unique in  $Y'$ . Let us show that  $\text{MinDeg}_i(Y) \leq s$ . By definition there are  $r_1, \dots, r_n \in \{0, 1, \dots, s-1\}$  such that  $(y'_1, \dots, y'_n) = (sx'_1 + r_1, \dots, sx'_n + r_n) \in Y$ . It is easy to see that  $(y'_1, \dots, y'_{i-1}, y'_{i+1}, \dots, y'_n) \in Y_{[n]/\{i\}}$  is connected with at most  $s$  left vertices of  $G_i(Y)$ . More precisely, the only possible neighbors of  $(y'_1, \dots, y'_{i-1}, y'_{i+1}, \dots, y'_n)$  are

$$sx'_i, sx'_i + 1, \dots, sx'_i + s - 1.$$

Indeed, otherwise there is  $x_i \neq x'_i$  such that  $(x'_1, \dots, x'_{i-1}, x_i, x'_{i+1}, \dots, x'_n) \in Y$ . The latter contradicts the fact that  $x'$  is  $i$ -unique in  $Y'$ .  $\blacktriangleleft$

**Proof of Theorem 9.** Due to Lemma 15 there is a reducible  $X' \subset \{0, 1, \dots, m-1\}^n$  such that for every  $i \in [n]$  we have  $\text{AvgDeg}_i(X') \geq n - \varepsilon$ . By Lemma 17, applied to  $X = \text{In}(X', s)$  for every  $i \in [n]$  we have:  $\text{AvgDeg}_i(X) \geq s(n - \varepsilon)$ . Finally, due to Lemma 18, for all non-empty  $Y \subset X$  there is  $i \in [n]$  such that  $\text{MinDeg}_i(Y) \leq s$ .  $\blacktriangleleft$

## 6 Expanders Similar to $AP_q$

In this section we prove Proposition 5. Let us stress that this Proposition is just a slight improvement of Proposition 10 for sets of size 2. Proposition 10 itself is not strong enough to conclude that in all  $(m^2, m, 1/\sqrt{m})$ -spectral expanders any two distinct vertices have at most 1 common neighbor.

For  $S \subset V$  let  $\mathbb{1}_S \in \mathbb{R}^{|V|}$  denote characteristic vector of a set  $S$ . Assume for contradiction that there are distinct  $u, v \in V$  such that  $|\Gamma(u) \cap \Gamma(v)| \geq 2$ . Then the size of  $\Gamma(\{u, v\})$  is at most  $2d - 2$ . Assume that  $M$  is the adjacency matrix of  $G$ . Denote  $w = \{u, v\}$ . Let us show that

$$\|M\mathbb{1}_w\|^2 \leq d^2 \left( 2\gamma^2 + \frac{4(1-\gamma^2)}{m} \right). \quad (5)$$

Indeed, observe that  $\mathbb{1}_w = \frac{2}{m}\mathbb{1}_V + (\mathbb{1}_w - \frac{2}{m}\mathbb{1}_V)$  and  $(\mathbb{1}_w - \frac{2}{m}\mathbb{1}_V)$  is perpendicular to  $\mathbb{1}_V$ . Since

#### 4:14 From Expanders to Hitting Distributions and Simulation Theorems

$G$  is a  $(m, d, \gamma)$ -spectral expander, this implies that

$$\begin{aligned} \|M\mathbb{I}_w\|^2 &= \left\| M \left( \frac{2}{m} \mathbb{I}_w \right) \right\|^2 + \left\| M \left( \mathbb{I}_w - \frac{2}{m} \mathbb{I}_V \right) \right\|^2 \leq \frac{4d^2}{m} + \gamma^2 d^2 \left\| \left( \mathbb{I}_w - \frac{2}{m} \mathbb{I}_V \right) \right\|^2 \\ &= \frac{4d^2}{m} + \gamma^2 d^2 \left( 2 \left( 1 - \frac{2}{m} \right)^2 + (m-2) \frac{4}{m^2} \right) \\ &= \frac{4d^2}{m} + \gamma^2 d^2 \left( 2 - \frac{4}{m} \right) = d^2 \left( 2\gamma^2 + \frac{4(1-\gamma^2)}{m} \right), \end{aligned}$$

and thus (5) is proved.

To obtain a contradiction it is enough to show the following inequality

$$\|M\mathbb{I}_w\|^2 \geq 2d + 4. \quad (6)$$

Assume that there are  $t \leq 2d - 2$  non-zero coordinates in  $M\mathbb{I}_w$ . Let  $\xi_1, \dots, \xi_t$  be the values of these coordinates. Their sum is  $2d$ . We need to show that  $\xi_1^2 + \dots + \xi_t^2 \geq 2d + 4$ . Observe that  $\xi_1 - 1, \dots, \xi_t - 1$  are non-negative integers and their sum is  $2d - t \geq 2$ . Clearly this implies that  $(\xi_1 - 1)^2 + \dots + (\xi_t - 1)^2 \geq 2$ . Indeed, otherwise the sum of  $\xi_1 - 1, \dots, \xi_t - 1$  is either 0 or 1. Hence

$$\xi_1^2 + \dots + \xi_t^2 = (\xi_1 - 1)^2 + \dots + (\xi_t - 1)^2 + 4d - t \geq 2 + 4d - t \geq 2d + 4.$$

### 7 Proof of Proposition 6

Denote  $s = 2^{k-h}$ . Assume that there is a 0-monochromatic rectangle  $A \times B$  of  $g$  such that  $|A| \geq s$  and  $B \geq s$ . Then clearly the proposition is true for  $b = 1$  and  $X = A, Y = B$ .

Now assume that if  $A \times B$  is a 0-monochromatic rectangle of  $g$ , then either  $|A| < s$  or  $B < s$ . Take  $\mathcal{X}, \mathcal{Y}$  independently and uniformly at random from the set of all  $s$ -element subsets of  $\{0, 1\}^k$ . Fix any 0-monochromatic rectangle  $A \times B$  of  $g$ . Let us show that  $\mathcal{X} \times \mathcal{Y}$  intersects  $A \times B$  with probability at most  $2^{k-2h+1}$ . Indeed, assume WLOG that  $|A| < s$ . Then

$$\Pr[\mathcal{X} \times \mathcal{Y} \cap A \times B \neq \emptyset] \leq \Pr[\mathcal{X} \cap A \neq \emptyset] = 1 - \frac{\binom{2^k - |A|}{s}}{\binom{2^k}{s}} \leq 1 - \frac{\binom{2^k - s}{s}}{\binom{2^k}{s}}$$

Since  $h \geq 1$ , we have that  $s \leq 2^k/2$ . Applying Lemma 13 we obtain:

$$\Pr[\mathcal{X} \times \mathcal{Y} \cap A \times B \neq \emptyset] \leq \frac{s^2}{2^k - s} \leq \frac{s^2}{2^k/2} = 2^{k-2h+1}.$$

Due to the standard averaging argument this means that for any probability distribution  $\mu$  over 0-monochromatic rectangles of  $g$  it is possible to fix  $\mathcal{X} = X, \mathcal{Y} = Y$  in such a way that

$$\Pr_{R \sim \mu} [R \cap X \times Y \neq \emptyset] \leq 2^{k-2h+1}.$$

---

**References**

---

- 1 Arkadev Chattopadhyay, Michal Koucký, Bruno Loff, and Sagnik Mukhopadhyay. Simulation theorems via pseudorandom properties. *arXiv preprint arXiv:1704.06807*, 2017.
- 2 Susanna F de Rezende, Jakob Nordström, and Marc Vinyals. How limited interaction hinders real communication (and what it means for proof and circuit complexity). In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pages 295–304. IEEE, 2016.
- 3 Mika Goos, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. *SIAM Journal on Computing*, 45(5):1835–1869, 2016.
- 4 Mika Göös, Toniann Pitassi, and Thomas Watson. Deterministic communication vs. partition number. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 1077–1088. IEEE, 2015.
- 5 Mika Göös, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for bpp. *arXiv preprint arXiv:1703.07666*, 2017.
- 6 Hamed Hatami, Kaave Hosseini, and Shachar Lovett. Structure of protocols for xor functions. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pages 282–288. IEEE, 2016.
- 7 Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 2006.
- 8 Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.
- 9 Ran Raz and Pierre McKenzie. Separation of the monotone nc hierarchy. In *Foundations of Computer Science, 1997. Proceedings., 38th Annual Symposium on*, pages 234–243. IEEE, 1997.
- 10 Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Annals of mathematics*, pages 157–187, 2002.
- 11 Salil P Vadhan et al. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, 2012.
- 12 Xiaodi Wu, Penghui Yao, and Henry Yuen. Raz-mckenzie simulation with the inner product gadget. In *Electronic Colloquium on Computational Complexity (ECCC)*, 2017.