Partition Expanders*

Dmitry Gavinsky and Pavel Pudlák

Institute of Mathematics, Academy of Sciences, Prague, Czech Republic
dmitry.gavinsky@gmail.com, pudlak@math.cas.cz

Abstract

We introduce a new concept, which we call *partition expanders*. The basic idea is to study quantitative properties of graphs in a slightly different way than it is in the standard definition of expanders. While in the definition of expanders it is required that the number of edges between any pair of sufficiently large sets is close to the expected number, we consider *partitions* and require this condition only for *most of the pairs* of blocks. As a result, the blocks can be substantially smaller.

We show that for some range of parameters, to be a partition expander a random graph needs *exponentially smaller* degree than any expander would require in order to achieve similar expanding properties.

We apply the concept of partition expanders in communication complexity. First, we give a PRG for the SMP model of the optimal seed length, $n + O\log k$. Second, we compare the model of SMP to that of Simultaneous Two-Way Communication, and give a new separation that is stronger both qualitatively and quantitatively than the previously known ones.

1998 ACM Subject Classification F.1.1 Models of Computation

Keywords and phrases partitions, expanders, communication, complexity

Digital Object Identifier 10.4230/LIPIcs.STACS.2014.325

1 Introduction

Expanders are a very interesting and useful concept and appear in many applications in computer science. Therefore several related concepts have been introduced; e.g., lossless expanders [6], monotone expanders and dimension expanders [7], superexpanders [13].

In this paper we introduce yet another concept that we call *partition expanders*. The definition is motivated by the following observation. The well-known Expander-Mixing Lemma says, roughly speaking, that for every two sufficiently big sets of vertices A and B the number of edges of the expander between A and B is close to $\frac{d}{n} \cdot |A| \cdot |B|$, where n is the number of vertices and d is the degree. If we want to apply this lemma to smaller sets, we have to increase the degree of expanders appropriately.

Now suppose we have a partition of the vertices of the graph and we only want to satisfy the density condition for most of the pairs of sets. It turns out that a random graph with relatively small degree is able to satisfy this condition for partitions with many blocks, although the Expander-Mixing Lemma is not able to give any interesting estimate. So while expanders are graphs with "typical connectivity" with respect to *subsets of vertices*, partition expanders have "typical connectivity" with respect to *partitions of vertices*. Informally speaking, in the context of expanders, partitions are "more structured" objects than subsets, and therefore demanding the same "expanding performance" with respect to partitions can be

* Partially funded by the grant P202/12/G061 of GA ČR and by RVO: 67985840.

31st Symposium on Theoretical Aspects of Computer Science (STACS'14).
Editors: Ernst W. Mayr and Natacha Portier; pp. 325–336

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

viewed as a relaxation of usual expanders. In return, we expect partition expanders to have considerably smaller degree than usual expanders with the same expanding performance.

There are several possible ways to formally define a partition expander. We choose the following definition as "canonical" due to its brevity and robustness. We will give alternative definitions shortly.

▶ **Definition 1.** Partition expanders Let $G = (V, E)$ be an (undirected) graph. Let μ be the uniform distribution over $V \times V$, and let μ_G be the uniform distribution over E. For any coloring $c : V \to [K]$, let ν^c and ν^c_G be the distributions of the pair $(c(v_1), c(v_2))$ when (v_1, v_2) is chosen according to μ or μ_G, respectively.

For $K \in \mathbb{N}$ and $\delta \in (0, 1)$, we say that G is a (K, δ)-*partition expander* if for every coloring $c : V \to [K]$ the statistical distance between ν^c and ν^c_G is at most δ.

It should be noted that this concept is interesting in the situations where the number K of partitions is increasing with the number of vertices and the graphs are d-regular with d increasing. We are mainly interested in the question of how small d can be for a given K, assuming $0 < \delta < 1$ is a fixed constant.

1.1 Our results

We start by giving several equivalent definitions of partition expanders, which emphasize the fact that they are a natural modification of usual expanders.

In Section 3 we analyze the behavior of random graphs as partition expanders. We prove that random d-regular graphs almost always are good partition expanders – the dependence of K on d is the best possible, namely exponential.

In Section 4 the notion of partition expanders is advocated through comparing it to expanders. We show that the gap between the absolute values of the first two eigenvalues does not ensure that the graph is a good partition expander. Namely, if only the spectral gap is taken into account when a partition expander is constructed, then the degree has to be *exponentially larger* than an optimal partition expander requires. Since the spectral gap characterizes almost tightly the expander properties of a graph, this demonstrates exponential advantage of partition expanders (in those scenarios when they are suitable) over expanders. In other words, if "partition expansion" is the desired behavior, then using an expander instead of an optimal partition expander would incur exponential loss in terms of the required degree.

Based on the spectral properties only, we use the Hoffman-Wielandt inequality and get a slightly better bound than what would follow from a direct application of the Expander-Mixing Lemma.[1] The fact that the spectral gap is incapable to characterize good partition expanders partially explains why new methods are required for their construction.

In Section 5 we present another equivalent definition of partition expanders. We show that a graph $G = (V, E)$ is a partition expander if and only if the uniform distribution over E is a *Pseudo-Random Generator (PRG)* in the setting of *Simultaneous Message Passing (SMP)* in communication complexity. We use this fact to give a lower bound on the degree of partition expanders, thus showing *optimality* of the randomized construction given in Section 3.

In the second part of Section 5 we show two applications of our randomized construction of a partition expander. First, we construct a PRG against SMP protocols of communication

[1] We get quadratic improvement in terms of the partition size, and show that it is essentially optimal general bound in terms of the spectral gap alone.

cost k that requires seed length $n + O(\log k)$ (see Theorem 15 and the comment thereafter).[2] Second, we compare the model of SMP to that of Simultaneous Two-Way Communication, and give a new separation that is stronger both qualitatively and quantitatively than the previously known ones (see Theorem 19).

2 Notation and more

Unless stated otherwise, all sets are assumed to be finite, and all graphs are *undirected* and *simple* (having no self loops or multiple edges).[3] For two subsets $S_1, S_2 \subseteq V$, we denote by $E(S_1, S_2)$ the set of ordered pairs (v_1, v_2) such that (v_1, v_2) is an edge in E, $v_1 \in S_1$ and $v_2 \in S_2$, and write $E(v_1, v_2)$ for $E(\{v_1\}, \{v_2\})$.[4] We will say that a set family $\sigma = \{C_1, \ldots, C_K\}$ is a K-partition of a set X if $\cup_{i=1}^{K} C_i = X$ and $C_1, \ldots, C_K$ are pairwise disjoint and nonempty.

The *statistical distance* between two distributions μ_1 and μ_2 defined over a set X is

$$\mathrm{d}_{\mathrm{st}}(\mu_1, \mu_2) \stackrel{\mathrm{def}}{=} \frac{1}{2} \sum_{x \in X} |\mu_1(x) - \mu_2(x)| .$$

▶ **Lemma 2.** *Let $K \in \mathbb{N}$ and $\delta \in \mathbb{R}$. The following statements are equivalent:*

1. *$G = (V, E)$ is a (K, δ)-partition expander.*
2. *For every K-partition $\sigma = \{C_1, \ldots, C_K\}$ of V,*

$$\delta \geq \frac{1}{2} \sum_{i,j \in [K]} \left| \frac{|E(C_i, C_j)|}{|E|} - \frac{|C_i| \cdot |C_j|}{|V|^2} \right| . \tag{1}$$

3. *For every K-partition σ and $S \subseteq [K] \times [K]$,*

$$\delta \geq \sum_{(i,j) \in S} \left(\frac{|E(C_i, C_j)|}{|E|} - \frac{|C_i| \cdot |C_j|}{|V|^2} \right) = \frac{\sum_S |E(C_i, C_j)|}{|E|} - \frac{\sum_S |C_i| \cdot |C_j|}{|V|^2} . \tag{2}$$

4. *Like 0c, but only over symmetric S (i.e., $(i, j) \in S \Leftrightarrow (j, i) \in S$).*

Proof. Equivalence between 0a and 0b is immediate from Definition 1. Equivalence between 0c and 0d follows from the fact that G is undirected. To see that 0b is equivalent to 0c, note that

$$\sum_{i,j \in [K]} \left(\frac{|E(C_i, C_j)|}{|E|} - \frac{|C_i| \cdot |C_j|}{|V|^2} \right) = \frac{\sum_{[K] \times [K]} |E(C_i, C_j)|}{|E|} - \frac{\sum_{[K] \times [K]} |C_i| \cdot |C_j|}{|V|^2} = 0.$$

◀

There are many possible ways to define expanders. The standard definition is based on the second largest absolute value of an eigenvalue of a graph G, which we will denote by $\lambda(G)$.

[2] All previously known PRGs in communication complexity were given against stronger models, thus requiring exponentially larger "overhead" over n in terms of seed length – for details, see Section 5.

[3] In those cases when we explicitly allow multiple edges, the edges of a graph will be viewed as a collection with repetitions.

[4] Note that if $v_1, v_2 \in S_1 \cap S_2$, then the edge (v_1, v_2) appears in $E(S_1, S_2)$ twice: as ordered pairs (v_1, v_2) and (v_2, v_1).

▶ **Definition 3** (Expanders). A regular graph G is an ℓ-expander if $\lambda(G) \leq \ell$.

We will denote the degree of a regular graph G by $d(G)$, or simply by d when G is clear from the context.

The most natural relation between expanders and partition expanders comes from the following well-known fact (e.g., see [2]).

▶ **Lemma 4** (Expander-Mixing Lemma). *Let (V, E) be an ℓ-expander. Then for every $S_1, S_2 \subseteq V$,*

$$\left| \frac{|E(S_1,S_2)|}{|E|} - \frac{|S_1| \cdot |S_2|}{|V|^2} \right| \leq \ell \cdot \frac{\sqrt{|S_1| \cdot |S_2|}}{|E|} = \frac{\ell}{d} \cdot \frac{\sqrt{|S_1| \cdot |S_2|}}{|V|}.$$

One can show using this lemma that an ℓ-expander is a (K, δ)-partition expander for constant $\delta > 0$ and certain $K \in \Theta(d/\ell)$ – however, this trivial arguments fails for $K \geq d/\ell$. In Section 4 we will use the Hoffman-Wielandt inequality to show that an ℓ-expander is a $(K, \Omega(1))$-partition expander for certain $K \in \Theta\big((d/\ell)^2\big)$, and that will be shown to be optimal up to the factor of $\log n$.

▶ **Theorem 5** (Hoffman-Wielandt inequality [9]). *If A and B are normal matrices with respective eigenvalues $\lambda_1(A), \ldots, \lambda_n(A)$ and $\lambda_1(B), \ldots, \lambda_n(B)$, then*

$$\min_{\pi} \left\{ \sum_{i=1}^{n} \left| \lambda_i(A) - \lambda_{\pi(i)}(B) \right|^2 \right\} \leq \|A - B\|_F^2 ,$$

where π runs over all permutations over $[n]$ and $\| \ldots \|_F^2$ denotes the square of the Frobenius norm (the sum of squares of the absolute values of the elements).

If A and B are symmetric real matrices, we can drop the absolute value and write the terms as $\lambda_i(A)^2 + \lambda_{\pi(i)}(B)^2 - 2\lambda_i(A)\lambda_{\pi(i)}(B)$. Since the sum of the squares of eigenvalues of a matrix is the square of its Frobenius norm, the inequality is equivalent to

$$\sum_{i,j} a_{ij} b_{ij} \leq \max_{\pi} \left\{ \sum_{i=1}^{n} \lambda_i(A) \lambda_{\pi(i)}(B) \right\}. \tag{3}$$

Let $d, n \in \mathbb{N}$ be such that $2|dn$, denote by $\mathcal{G}_{n,d}$ the uniform distribution on d-regular (simple undirected) graphs on n vertices. In our analysis we will use the *pairing method* for generating $G \sim \mathcal{G}_{n,d}$, due to Bollobás [5] (also see [14]).

▶ **Lemma 6** (Pairing method [5]). *The following procedure generates $E \subseteq [n] \times [n]$ such that $G = ([n], E) \sim \mathcal{G}_{n,d}$.*

1. *Let $\pi \subset [nd] \times [nd]$ be a uniformly random perfect matching on $[nd]$ (viewed as a symmetric set of directed edges). For $i \in [n]$, let $cell_i \stackrel{def}{=} \{x \mid id - d < x \leq id\}$ and $d_\pi(v_1, v_2) \stackrel{def}{=} |\pi(cell_{v_1}, cell_{v_2})|$.*
2. *For every $(v_1, v_2) \in [n] \times [n]$, let (v_1, v_2) be $d_\pi(v_1, v_2)$ times an element of E.*
3. *Return to Step 0a if $G = ([n], E)$ is not simple.*

In the analysis we will consider the distribution of $([n], E)$ resulting from dropping Step 0c off the above procedure; let us denote it by $\mathcal{G}'_{n,d}$. Observe that a graph $G \sim \mathcal{G}'_{n,d}$ is always undirected, but doesn't have to be simple.[5]

We will use the following estimate, due to McKay and Wormald [12]:

[5] Note also that the distribution $\mathcal{G}'_{n,d}$ is not uniform over its support - e.g., $\mathcal{G}'_{2,2}$ produces the graph with two parallel edges with probability 2/3.

▶ **Lemma 7** ([12]). *For $d \in o(\sqrt{n})$,*

$$\Pr_{G\sim\mathcal{G}'_{n,d}}[G \text{ is simple}] \in \exp\left(\frac{1-d^2}{4} - \frac{d^3}{12n} + O\left(\frac{d^2}{n}\right)\right) \subseteq \exp(o(n)).$$

3 Random d-regular graphs as partition expanders

Let us see that a random regular graph is likely to form a partition expander.

▶ **Theorem 8.** *For $d \in O\left(n^{1/3}\right)$, a random d-regular simple undirected graph on n vertices is a (K,δ)-partition expander with probability at least $1 - \exp\left(n\log K + K^2 - \Omega\left(\delta^2 nd\right)\right)$.*

The proof can be found in the full version of the paper.

▶ **Corollary 9.** *For any $\varepsilon > 0$ and $B \in \mathbb{N}$ there exists $C \in \mathbb{N}$, such that the following holds: A random d-regular graph on n vertices is a (K,δ)-partition expander with probability at least $1-\varepsilon$, as long as $K \le B\cdot\sqrt{n}$ and $d \ge \frac{C\cdot\log K}{\delta^2}$.*

4 Partition expanders vs. expanders

Let us compare the notions of expanders and partition expanders in more detail.

▶ **Theorem 10.** *Let G be a d-regular ℓ-expander on n vertices. Then it is a $(K, \sqrt{K}\ell/d)$-partition expander for every $K < d^2/\ell^2$.*

The proof is based on the Hoffman-Wielandt inequality and can be found in the full version of the paper.

Note that the Expander-Mixing Lemma (Lemma 4) only gives that G is a (K,δ) partition expander for $\delta = O(K\ell/d)$, which is meaningful only for $K < d/\ell$. The statement of the above theorem is essentially tight (cf. Theorem 12), and this means that only small (quadratic, in terms of K vs. d) improvement can result from using partition expanders instead of expanders, as long as the construction of a partition expanders relies on the spectral gap. On the other hand, we will see soon that good partition expanders offer *exponential* improvement in terms of the dependence of K on d.

Now we will show that the above bound is essentially optimal, and therefore, in general expanders are not good partition expanders. We will use the following result of Alon and Roichman [1]. (For a simpler proof, and an explicit and better bound, see [11].)

▶ **Theorem 11** ([1, 11]). *There exists an absolute constant c such that for every finite group Γ and any $d \le |\Gamma|$, the following is true. If we pick uniformly at random the elements $g_1,\ldots,g_d \in \Gamma$, then the resulting Cayley-graph has the second largest eigenvalue λ satisfying*

$$\lambda \le c\cdot\sqrt{d\log|\Gamma|}$$

with probability going to 1 as $|\Gamma| \to \infty$.

This theorem is not stated explicitly in those papers, but it is an immediate corollary of Theorem 2 of [11]. (One can take any constant c such that $c > 2\ln 2$.)

Let $m > 0$ be a natural number and let Γ be the symmetric group on m elements represented by permutations of $[m]$. Let $\pi_1,\ldots,\pi_d$ be some permutations for which the bound on the eigenvalue is satisfied. W.l.o.g. we will assume that for every $i \in [d]$ there is a $j \in [d]$ such that $\pi_j = \pi_i^{-1}$. Let G be the Cayley graph determined by Γ and $\pi_1,\ldots,\pi_d$.

Let $1 \leq t \leq m$. We will consider the partition $\{C_1, \ldots, C_K\}$ defined by the following equivalence relation on G

$$\rho|_{[t]} = \sigma|_{[t]},$$

where $\rho, \sigma \in G$ are permutations and $|_{[t]}$ denote their restriction to the first t elements. Thus the number of blocks is $K = m(m-1)\ldots(m-t+1)$. Consider the symmetric set S defined by

$$(i,j) \in S \;\equiv\; \exists \rho \in C_i, \sigma \in C_j \; \exists s \in [d] \; \rho|_{[t]} = \pi_s \sigma|_{[t]} \,. \tag{4}$$

Note that if for some i and j the condition is satisfied by some $s = s_0$, then for all $\rho \in C_i, \sigma \in C_j$, we have $\rho|_{[t]} = \pi_{s_0}\sigma|_{[t]}$.

Consider the equation (2) that defines partition expanders. The first term is in our case equal to 1. To bound the second term, note that for a given $s \in [d]$ the number of pairs ρ, σ satisfying the condition $\rho|_{[t]} = \pi_\ell \sigma|_{[t]}$ is $m!(m-t)!$. Hence the second term is bounded by

$$\frac{d \cdot m!(m-t)!}{(m!)^2} = \frac{d}{m(m-1)\ldots(m-t+1)} = \frac{d}{K}.$$

This proves that if $d/K < 1-\delta$, then G is not a (K,δ)-partition expander.

Thus we have proved:

▶ **Theorem 12.** *There exist a constant c such that for infinitely many n and every $d \leq n$, there are d-regular $c\sqrt{d \log n}$-expanders on n vertices which are not $(K, 1 - \frac{d+1}{K})$-partition expanders.*

Comparing this statement to the bound given by Theorem 10 in the most natural regime when a $(K, 1-\Omega(1))$-partition expander is required, we can see that the upper and the lower bounds match up to the factor of $\log n$ in the spectral gap. In particular, since the second eigenvalue of a graph is always $\Omega\left(\sqrt{d}\right)$, K can be at most linear in d, as long as our only assumption about G is the absolute value of its second eigenvalue. In contrast to this, according to Corollary 9, there exist $(K, 1-\Omega(1))$ partition expanders whose degree is $O(\log K)$. Thus any construction of such partition expanders must rely on some properties of G, other than the spectral gap.

5 Partition expanders as PRGs in communication complexity

Let us turn to the realm of communication complexity, where we give an equivalent formulation of partition expanders. First, we use this equivalence to give a nearly-tight lower bound on the degree of good partition expanders, thus arguing near-optimality of the randomized construction given in Section 3. Second, we use the same construction to obtain a new separation between two models of communication complexity, which is qualitatively stronger than the previously known one.

We will use the following models of two-party communication complexity.

▶ **Definition 13** (Models of communication complexity)**.** Two players whose names are Alice and Bob each receive a binary string of length n, respectively denoted by x and y. Players' goal is to compute the value of $f(x,y)$, where $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is fixed. The players obey the following scenario:

- In the model of *Simultaneous Message Passing* (SMP), denoted by $\mathcal{R}^{\|}$, both Alice and Bob send a message to the third participant, the referee. The referee does not know the values of x and y, so his only input are the messages received from the players, and he has to produce the answer using the information received from the players. All three participants are allowed to use private randomness.
- The model of *SMP with shared randomness,* denoted by $\mathcal{R}^{\|,pub}$, is similar to $\mathcal{R}^{\|}$ but the players are allowed to use public randomness.[6]
- In the model of *One-Way Communication,* denoted by $\mathcal{R}^{1}$, Alice sends her message to Bob, who has to produce the answer using his part of the input and the information received from Alice.
- In the model of *Simultaneous Two-Way Communication,* denoted by $\mathcal{R}^{\leftrightarrow}$, Alice and Bob send their messages simultaneously, similarly to the case of SMP. But here the recipient of Alice's message is Bob, and the recipient of Bob's message is Alice. Upon receiving the partner's message, each player must produce an answer.

We say that a communication protocol solves the problem represented by f if it produces the correct answer(s) with probability at least $2/3$ for every possible input. The communication cost of a protocol is the maximal total number of bits sent by the players, and the communication cost of a function f is the minimal communication cost of a protocol that solves it in the given model.

The models $\mathcal{R}^{\|}$, $\mathcal{R}^{\|,pub}$ and $\mathcal{R}^{1}$ have been studied widely and the corresponding notation is commonly used; the Simultaneous Two-Way model has been considered in several works (see below), but no specific name was assigned to it. Note that when we say that an $\mathcal{R}^{\leftrightarrow}$-protocol has produced the answer "a", we refer to the situation when both the players have produced the same answer.

▶ **Definition 14** (Pseudo-randomness in communication complexity)**.** Let $\mathcal{M}$ be a communication complexity model, and let μ be a distribution defined over $\{0,1\}^n \times \{0,1\}^n$. We say that μ is k-pseudo-random for $\mathcal{M}$ if for any protocol $\mathcal{P}$ of communication cost at most k it holds that

$$\mathop{\mathbf{Pr}}_{(X,Y)\sim\mu}[\mathcal{P}(X,Y) \text{ outputs "1"}] - \mathop{\mathbf{Pr}}_{(X,Y)\sim\mathcal{U}_{\{0,1\}^n\times\{0,1\}^n}}[\mathcal{P}(X,Y) \text{ outputs "1"}] < \frac{1}{3}.$$

We say that $g : \{0,1\}^s \to \{0,1\}^n \times \{0,1\}^n$ is a k-Pseudo-Random Generator (k-PRG) of seed length s against $\mathcal{M}$ if the distribution of $g(X)$ when $X \sim \mathcal{U}_{\{0,1\}^s}$ is k-pseudo-random for $\mathcal{M}$.

Pseudo-randomness in the context of communication complexity has been introduced in [10]. Intuitively, both pseudo-randomness and lower bounds on communication cost can be viewed as claims that certain problem is hard for the model under consideration.

Given a d-regular graph $G = (\{0,1\}^n, E)$, let μ_G be the uniformly random distribution of the edges from E. Note that in order to choose $(v_1, v_2) \sim \mu_G$, a "seed" of length $n + \log d$ is both necessary and sufficient.

▶ **Theorem 15.** *Let $k, n \in \mathbb{N}$ and $G = (\{0,1\}^n, E)$. The following statements are equivalent:*

1. *G is a $(2^{\Theta(k)}, \delta)$-partition expander for some $\delta < 1/3$.*
2. *μ_G is $\Theta(k)$-pseudo-random for $\mathcal{R}^{\|}$.*

[6] Note that in the communication complexity setting Alice and Bob collaborate, and therefore availability of public randomness is equivalent to players' ability to use mixed strategies.

In particular, our construction in Section 3 corresponds to a k-PRG against $\mathcal{R}^{\|}$ of seed length $n+O(\log k)$. Note that due to the fact that in the context of communication complexity the players are computationally unlimited, a randomized construction of a PRG is neither meaningless nor trivial.[7]

Proof. Let C be a constant. First, suppose that G is a $(2^{Ck}, \delta)$-partition expander. Let $\mathcal{P}$ be an $\mathcal{R}^{\|}$-protocol of cost at most Ck, and let us show that it cannot distinguish with high confidence μ_G from $\mathcal{U}_{\{0,1\}^n \times \{0,1\}^n}$. Without loss of generality assume that $\mathcal{P}$ is deterministic, and let $\alpha : \{0,1\}^n \to \{0,1\}^{Ck}$ be the mapping from x to the concatenation of Alice's and Bob's messages in response to the input (x, x). Let $\nu_{\mathcal{U}}$ and ν_G be the distributions of $(\alpha(X), \alpha(Y))$ when, respectively, $(X,Y) \sim \mathcal{U}_{\{0,1\}^n \times \{0,1\}^n}$ and $(X,Y) \sim \mu_G$. Clearly,

$$\Pr_{(X,Y)\sim\mu_G}[\mathcal{P}(X,Y) \text{ outputs "1"}] - \Pr_{(X,Y)\sim\mathcal{U}_{\{0,1\}^n\times\{0,1\}^n}}[\mathcal{P}(X,Y) \text{ outputs "1"}] \leq \mathrm{d}_{\mathrm{st}}(\nu_G, \nu_{\mathcal{U}}).$$

Note that α defines a partition of $\{0,1\}^n$ into at most 2^{Ck} blocks, and by the definition of partition expanders,

$$\mathrm{d}_{\mathrm{st}}(\nu_G, \nu_{\mathcal{U}}) \leq \delta < 1/3.$$

Therefore, μ_G "fools" $\mathcal{P}$ and thus it is Ck-pseudo-random for $\mathcal{R}^{\|}$.

Now assume that μ_G is $2Ck$-pseudo-random for $\mathcal{R}^{\|}$, and let us show that G is a partition expander. Let $\sigma = \{S_1, \ldots, S_{2^{Ck}}\}$ be a partition of $\{0,1\}^n$, and for $x \in \{0,1\}^n$, define $\sigma(x) \stackrel{\text{def}}{=} i$ for i such that $x \in S_i$. Let $\mathcal{P}_\sigma$ be an $\mathcal{R}^{\|}$-protocol, where upon receiving input (X,Y), Alice sends $\sigma(X)$ and Bob sends $\sigma(Y)$. Let $\tau_{\mathcal{U}}$ and τ_G be the distributions of $(\sigma(X), \sigma(Y))$ when, respectively, $(X,Y) \sim \mathcal{U}_{\{0,1\}^n \times \{0,1\}^n}$ and $(X,Y) \sim \mu_G$. Note that $\mathcal{P}_\sigma$ is of cost $2Ck$, and therefore

$$\mathrm{d}_{\mathrm{st}}(\tau_{\mathcal{U}}, \tau_G) < 1/3,$$

since otherwise the referee would be able to distinguish the two cases with confidence high enough to contradict pseudo-randomness of μ_G. Let δ be the maximum value of $\mathrm{d}_{\mathrm{st}}(\tau_{\mathcal{U}}, \tau_G)$ possible under any choice of 2^{Ck}-partition σ, then $\delta < 1/3$ and G is a $(2^{Ck}, \delta)$-partition expander, as required. ◀

5.1 Lower bound on the degree of partition expanders

Let us use the correspondence between partition expanders and pseudo-random generators given by Theorem 15 in order to get a lower bound on the degree of partition expanders.

▶ **Theorem 16.** *For any $\delta < 1/3$, if a d-regular graph G is a (K, δ)-partition expander then $d \in \Omega\left(\frac{\log K}{\log\log K}\right)$.*

In particular, the randomized construction given in Section 3 is optimal, up to the multiplicative $\log\log K$ factor.

[7] For example, the models $\mathcal{R}^{1}$ and $\mathcal{R}^{\leftrightarrow}$ (and more generally, any two-party model where a k-bit message from one player reaches the other player, who also receives his own n bits of input) require seed length at least $n + k - O(1)$ even with a non-uniform PRG, as witnessed by the protocol where the sender sends the first k bits of his input and the computationally-unlimited recipient outputs "1" only if the message together with his own n bits of input have Kolmogorov complexity $n + k - O(1)$.

Proof. For convenience, let n and d be powers of 2. Let $G = ([n], E)$, and assume it is a (K, δ)-partition expander. On the one hand, according to Theorem 15, μ_G is $\Omega(\log K)$-pseudo-random for the SMP model. On the other hand, we will see below that an SMP protocol of cost $O(d \log d)$ can distinguish μ_G from the uniform distribution with error at most $1/4$, and therefore $d \in \Omega\left(\frac{\log K}{\log \log K}\right)$, as required.

The distinguishing protocol is as follows. When her input is $v \in V$, Alice sends to the referee the first $\log d + 2$ bits of the indices of the d neighbors of v. On input $u \in V$, Bob sends to the referee the first $\log d + 2$ bits of the index of u. The referee guesses that the input pair (v, u) has been drawn from the distribution μ_G if the index-prefix received from Bob appears in the list of d index-prefixes received from Alice. This protocol is always correct if the input comes from the support of μ_G, and errs with probability at most $1/4$ when the input comes from the uniform distribution. ◀

5.2 Model separations based on PRGs

Model separation in computational complexity usually means demonstrating existence of a computational problem that can be solved efficiently in one model, but not in the other. If several classes of problems can be handled by the models under consideration, one can define the corresponding *types* of model separations. When one problem class is a special case of another, separation via an element of the smaller class can be viewed as a stronger indication that the compared models have different computational power than separation via an element of the bigger class. These ideas can be pushed further, resulting in various "hierarchies" of model separations.

In the case of communication complexity, there are at least four natural classes of computational problems[8], namely:

- Total functions $f : A \times B \to Z$
- Partial functions $f : C \to Z$, $C \subseteq A \times B$
- Relations $\mathcal{P} \subseteq A \times B \times Z$
- Distinguishing some distribution μ defined on $A \times B$ from the uniform (cf. Definition 14)

Consider the four types of model separations corresponding to these four classes. We will call the fourth type *separation via a PRG*. Obviously, if two communication models are separable via a total function they are also separated via a partial function, and separability via a partial function implies separability via a relation. On the other hand, there are pairs of communication models that can be separated via a relation but not via a partial function (e.g., see [8]), and there are many pairs of models that have been separated via partial functions, but are *conjectured* not to be separable via total functions (e.g., most of quantum communication models form such pairs with their natural classical counterparts). Therefore, in communication complexity it is always desirable to separate models via the "most limited" possible type of separation, as that gives the "strongest" possible indication of difference in the computational power of those models.

To the best of our knowledge, separation via a PRG has not been studied in the context of communication complexity. It is probably incomparable to the first three types of separation: On the one hand, it is straightforward to get a separation via a PRG by modifying slightly one of the known separations via a partial function between quantum and classical one-way

[8] The same applies to many other fields of complexity, where also most of the following discussion remains valid – e.g., in the field of circuit complexity.

models, but it is *conjectured* that those two models cannot be separated via a total function. Therefore, modulo that conjecture, separation via a PRG cannot, in general, be as limited as separation via a total function. On the other hand, the models $\mathcal{R}^{\|}$ and $\mathcal{R}^{\|,pub}$ cannot be separated via a PRG (in general, it is easy to see that for any distribution-distinguishing task there exists an optimal protocol that does not need any randomness), but they can be separated via a total function – e.g., the equality function. Therefore, separation via a total function cannot, in general, be as limited as separation via a PRG.

Is there a type of model separation that would be the most limited, and therefore separations demonstrated through it would be the most "convincing" indication of difference in the computational power of the compared models?

Take a total Boolean function $f : A \times B \to \{0,1\}$, let $\mathcal{M}$ be a communication complexity model, and consider the following two claims:

- No protocol in $\mathcal{M}$ of cost less than k can compute f.
- The distributions $\mathcal{U}_{f^{-1}(0)}$ and $\mathcal{U}_{f^{-1}(1)}$ are k-PRGs for $\mathcal{M}$.

We will say that f is *k-hard for $\mathcal{M}$* in the first case, and that f is *k-pseudo-random for $\mathcal{M}$* in the second.[9] If f is k-pseudo-random for $\mathcal{M}$, then it is also k-hard for $\mathcal{M}$; the converse is not necessarily true, as follows from the same example of the equality function in $\mathcal{R}^{\|}$.

As usual in communication complexity, we will say that a communication problem is *easy* for a given model if it can be solved by a protocol of cost $(\log n)^{O(1)}$.

▶ **Definition 17** (Ultra-separation)**.** Complexity models $\mathcal{M}_1$ and $\mathcal{M}_2$ are ultra-separated if there is a total Boolean function f that is easy for $\mathcal{M}_1$ and $n^{\Omega(1)}$-pseudo-random for $\mathcal{M}_2$.

Ultra-separation is a very limited type of model separation – in fact, the most limited "reasonable" one we came up with.

▶ Claim 18. For any two models that allow efficient error reduction for total functions, ultra-separability implies separability both via a total function and via a PRG.

Here by efficient error reduction we mean that if f can be solved efficiently, then for any constant ε there exists an efficient protocol that solves f with error at most ε. Probably all studied communication complexity models satisfy this very natural property.

Proof. If f is $n^{\Omega(1)}$-pseudo-random for $\mathcal{M}_2$, then it is also $n^{\Omega(1)}$-hard for $\mathcal{M}_2$, and therefore ultra-separability implies separability via a total function.

If f is easy for $\mathcal{M}_1$, then the elements of $f^{-1}(1)$ can be distinguished from the elements of $f^{-1}(0)$ with worst-case error at most $1/10$ by a protocol of cost $(\log n)^{O(1)}$. Without loss of generality, let $\mathbf{Pr}\left[f(X,Y)=1\right] \leq 1/2$ when (X,Y) is uniformly random. Then there exist an efficient protocol in $\mathcal{M}_1$ that outputs "1" with probability at least $9/10$ when $(X,Y) \sim \mathcal{U}_{f^{-1}(1)}$, and with probability at most $11/20$ when (X,Y) is uniformly random. So, $\mathcal{U}_{f^{-1}(1)}$ can be distinguished from the uniform with "bias" more than $1/3$ by an efficient protocol in $\mathcal{M}_1$, and thus it is not a PRG. Therefore, ultra-separability implies separability via a PRG. ◀

5.3 Ultra-separation of $\mathcal{R}^{\|,pub}$ and $\mathcal{R}^{\leftrightarrow}$

We have seen that *ultra-separability of two models is a stronger evidence of difference in their computational power than separability via a function (total or partial), via a relation, or via a PRG*. We are not aware of any type of model separation that would not be subsumed by

[9] Note that we required both $\mathcal{U}_{f^{-1}(0)}$ and $\mathcal{U}_{f^{-1}(1)}$ to be k-PRGs for $\mathcal{M}$ when f is k-pseudo-random in order not to require f to be balanced; if it is balanced, either condition implies the other.

ultra-separation. Therefore, it is interesting to demonstrate ultra-separations even for those pairs of models that have been separated previously via some "less convincing" methods.

For long time, it had been believed that the models $\mathcal{R}^{\|,pub}$ and $\mathcal{R}^{\leftrightarrow}$ were equivalent. In 2002 Bar-Yossef, Jayram, Kumar and Sivakumar [4] demonstrated a separation between these models via a cleverly constructed total function g, for which $\mathcal{R}^{\leftrightarrow}(g) \in O(\log n)$ and $\mathcal{R}^{\|,pub}(g) \in \Omega(\sqrt{n})$. The ideas used in their construction seem to be insufficient to yield separation via a PRG.

▶ **Theorem 19.** *The models $\mathcal{R}^{\|,pub}$ and $\mathcal{R}^{\leftrightarrow}$ can be ultra-separated. Namely, there exists a total Boolean function f, such that $\mathcal{R}^{\leftrightarrow}(f) \in O(\log n)$ and $\mathcal{U}_{f^{-1}(1)}$ cannot be distinguished from $\mathcal{U}_{\{0,1\}^n \times \{0,1\}^n}$ by any $\mathcal{R}^{\|,pub}$-protocol of cost $o(n)$.*

The new separation is stronger not only qualitatively, but quantitatively as well – the improvement results from the (optimal) lower bound of $\Omega(n)$ on the $\mathcal{R}^{\|,pub}$-complexity of f.

Proof. From Corollary 9 it follows that for any constant δ there exists a graph G on 2^n vertices of degree $d \in \Theta(n)$, which is a $(2^{n/2}, \delta)$-partition expander. According to Theorem 15, the corresponding μ_G is $\Theta(n)$-pseudo-random for $\mathcal{R}^{\|}$. Clearly, the same is true for $\mu_{\bar{G}}$, where $\bar{G}$ is the complement graph. If we define $f_G : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ to be the "edge function" of G, then it is $\Theta(n)$-pseudo random for $\mathcal{R}^{\|}$.

Let us see that $\mathcal{R}^{\leftrightarrow}(f_G) \in O(\log d) = O(\log n)$. Consider a protocol where the players use shared randomness[10] to choose a hash function from $\{0,1\}^n$ to $\{0,1\}^{2\log d}$, then Alice sends the hash-value of x and Bob answers "1" if the received value equals the hash-value of one of the neighbors of y in G, and "0" otherwise (if Bob is the sender, they act symmetrically). This protocol has communication cost $O(d)$ and computes f_G with error $o(1)$. The result follows. ◀

6 Discussion

The most interesting open problem is to find an explicit construction of a good partition expander; more precisely, to construct a family of (K, δ)-partition expanders in which $\delta < 1$ is constant, K goes to infinity, and the degrees are $d = O(\log K)$. We will call informally such families of graphs *good partition expanders*. As we have shown in this paper, expanders are, in general, not good partition expanders and it seems unlikely that the property would be implied by a property of the spectrum of a graph. One possible way of constructing good partition expanders could be by using zig-zag product or a similar kind of product. Indeed, in a recent paper [13] Mendel and Naor have shown that zig-zag product can be used for constructing various types of generalizations of expanders. These constructions start with a small object, which can be found by brute force, and which are enlarged by applying products repeatedly. They work well when one needs constant degree, but in our case we need increasing degree and to satisfy a certain property for partitions with exponentially increasing number of blocks. It is not totally excluded that some kind of product will work, but it will require a new kind of argument to prove it.

We demonstrated some applications of partition expanders in communication complexity. In particular, we defined the notion of ultra-separation and argued that it is one of the weakest model-separating methods, thus applying it provides a very strong (probably, the strongest known) evidence that the two separated models have different computational power.

[10] The power of the model $\mathcal{R}^{\leftrightarrow}$ is not affected by allowing public randomness.

We gave an example of such separation. It would be interesting to find more examples of ultra-separations, not only in communication complexity.

We believe that partition expanders will be useful in many other areas of complexity theory, especially when explicit constructions are found. For example, one could use good partition expanders instead of expanders in the pseudorandom generators of Impagliazzo, Nisan and Wigderson [10], provided that an explicit construction of good partition expanders is found. Since the number of partitions corresponds to the exponential of space complexity, they would certainly have better parameters. This, however, requires further research, because the direct application of partition expanders in INW generators does not seem to give substantially better results than the use of expanders.

Acknowledgments. We thank Hartmut Klauck and anonymous reviewers for helpful comments.

References

1 N. Alon and Y. Roichman. Random Cayley Graphs and Expanders. *Random Structures and Algorithms 5*, pages 271–284, 1994.

2 N. Alon and J. Spencer. The Probabilistic Method. *John Wiley*, 2008.

3 K. Azuma. Weighted Sums of Certain Dependent Random Variables. *Tohoku Mathematical Journal 68*, pages 357–367, 1967.

4 Z. Bar-Yossef, T. S. Jayram, R. Kumar, and D. Sivakumar. Information Theory Methods in Communication Complexity. *Proceedings of 17th IEEE Conference on Computational Complexity*, pages 93–102, 2002.

5 B. Bollobás. A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs. *European Journal of Combinatorics 1*, pages 311–316, 1980.

6 M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson. Randomness Conductors and Constant-Degree Lossless Expanders. *Proceedings of the 34th Symposium on Theory of Computing*, pages 659–668, 2002.

7 Z. Dvir and A. Wigderson. Monotone Expanders: Constructions and Applications. *Theory of Computing 6(1)*, pages 291–308, 2010.

8 D. Gavinsky, O. Regev, and R. de Wolf. Simultaneous Communication Protocols with Quantum and Classical Messages. *Chicago Journal of Theoretical Computer Science*, 7, 2008.

9 A. J. Hoffman and H. W. Wielandt. The Variation of the Spectrum of a Normal Matrix. *Duke Mathematical Journal 20*, pages 37–39, 1953.

10 R. Impagliazzo, N. Nisan, and A. Wigderson. Pseudorandomness for Network Algorithms. *Proceedings of the 26th Symposium on Theory of Computing*, pages 356–364, 1994.

11 Z. Landau and A. Russell. Random Cayley Graphs are Expanders: A Simple Proof of the Alon-Roichman Theorem. *Electronic Journal of Combinatorics 11*, 2004.

12 B. D. McKay and N. C. Wormald. Asymptotic Enumeration by Degree Sequence of Graphs with Degrees $o(n^{1/2})$. *Combinatorica 11(4)*, pages 369–382, 1991.

13 M. Mendel and A. Naor. Nonlinear Spectral Calculus and Super-Expanders. *Publications mathématiques de l'IHÉS*, 2013.

14 N. C. Wormald. Models of Random Regular Graphs. *Surveys in Combinatorics. Lecture Note Series 276*, pages 239–298, 1999.