Electronic Colloquium on Computational Complexity - Reports Series 1995 - available via:
ECCC
TR95-013
FTP: ftp.eccc.uni-trier.de:/pub/eccc/
WWW: http://www.eccc.uni-trier.de/eccc/
Email: ftpmail@ftp.eccc.uni-trier.de with subject 'help eccc'

The Parallel Repetition Conjecture for Trees is True

Oleg Verbitsky*
Department of Mechanics and Mathematics
Lviv State University
Universytetska 1, 290602 Lviv, Ukraine

Abstract

The parallel repetition conjecture (PRC) of Feige and Lovasz says that the error probability of a two prover one round interactive protocol repeated n times in parallel is exponentially small in n. We show that the PRC is true in the case when the bipartite graph of dependence between queries to provers is a tree. Previously this was known only in the case of complete bipartite graphs (i.e. when the queries to provers are independent). We suggest also the combinatorial characterization of method that was used to obtain most results towards the PRC and discuss some related combinatorial problems.

1 Introduction

A two prover one round interactive proof system is a both probabilistic and nondeterministic computational model for recognizing a language L. In this model, the polynomial time verifier separately sends messages to two computationally unlimited provers. After receiving provers' replies, the verifier must determine with low error probability whether or not an input w belongs to L.

*Supported in part by grant No. MGT 000 from the International Science Foundation and by an AMS-FSU grant.

On a fixed input w, a two prover one round interactive protocol is well described by a two person game G of the following type. Let X, Y, S, T be finite sets. Let ϕ be a predicate on $X \times Y \times S \times T$. A pair (x, y) is chosen randomly and uniformly from a set $Q \subseteq X \times Y$. The element x is revealed to Player *1*, the element y is revealed to Player *2*. Players *1* and *2* reply with $f(x) \in S$ and $h(y) \in T$ in accordance with their *strategies* $f : X \longrightarrow S$ and $h : Y \longrightarrow T$. If $\phi(x, y, f(x), h(y)) = 1$, then both players win; otherwise they lose. The objective of Players *1* and *2* is to maximize collectively the winning probability (taken over the uniform distribution of (x, y) on Q). The winning probability for the optimal players' strategies is denoted by $\omega(G)$. So,

$$\omega(G) = \max_{f,h} \mathbf{P}\left[\phi(x, y, f(x), h(y)) = 1\right].$$

We call the game G *nontrivial* if $\omega(G) \neq 1$.

We define an *n-product game* G^n with winning probability $\omega(G^n)$ as the execution of n independent copies of G in parallel. More formally, a collection $\langle (x_1, y_1), \ldots, (x_n, y_n) \rangle$ is chosen at random from Q^n. Players *1* and *2* each are supplied with n-vectors $\bar{x} = (x_1, \ldots, x_n)$ and $\bar{y} = (y_1, \ldots, y_n)$, and reply with n-vectors $F(\bar{x}) = (f_1(\bar{x}), \ldots, f_n(\bar{x}))$ and $H(\bar{y}) = (h_1(\bar{y}), \ldots, h_n(\bar{y}))$, respectively. Now the players win in the case $\bigwedge_{i=1}^{n} \phi(x_i, y_i, f_i(\bar{x}), h_i(\bar{y})) = 1$.

The only presently known relations between $\omega(G)$ and $\omega(G^n)$ are

$$(\omega(G))^n \leq \omega(G^n) \leq \omega(G).$$

The value $\omega(G)$ corresponds to the error probability of an interactive proof system. Fortnow, Rompel and Sipser [8] suggested to decrease the error probability by running the interactive protocol independently n times in parallel. To substantiate this approach we need a good upper bound on $\omega(G^n)$. The conjecture $\omega(G^n) = (\omega(G))^n$ was disproved by Fortnow [7] who presented the game G for which $\omega(G^2) > (\omega(G))^2$. Feige [5] improved this by giving an example of the nontrivial game G with $\omega(G^2) = \omega(G)$.

Denote $\bar{\omega}(G) = \sup_n (\omega(G^n))^{1/n}$. The following conjecture by Feige and Lovasz seems more realistic.

The Parallel Repetition Conjecture (further on PRC). *If G is nontrivial then $\bar{\omega}(G) < 1$.*

It will be sometimes convenient to consider the subset $Q \subseteq X \times Y$ as a bipartite graph with vertex classes X and Y. The PRC for the case $Q =$

$X \times Y$, i.e. if Q is a complete bipartite graph, was established by Cai, Condon and Lipton [2]. Their estimate on $\omega(G^n)$ was improved by Lapidot and Shamir [11] and Peleg [12] in the case $|X| = |Y| = 2$ and by Feige [5] and Alon [1] in the general case.

Some examples of nonfree games with exponentially decreasing $\omega(G^n)$ were used in [3]. Feige and Lovász [6] obtained the exponentially small upper bounds on $\omega(G^n)$ for the class of (nonfree) games with the *uniqueness* property defined in [2].

In this paper we examine one more case, namely, if a bipartite graph Q is a tree. We show that in this case the PRC also is correct, that is,

for any nontrivial game $G = \langle Q, S, T, \phi \rangle$, where Q is a tree, it holds $\bar{\omega}(G) < 1$.

The only fact known in the general case is the estimate $\omega(G^n) \leq r_k(n)$, $k = |Q|$, where $r_k(n)$ is the extremal density in Hales-Jewett theorem from Ramsey theory [14] (for more details see Section 2).

As the simple analysis shows, the methods of [2, 11, 5, 1, 14] as well as the method used here for trees start from one and the same basic idea that can be formalized as follows. Let us consider Q^n as a bipartite graph with vertex classes X^n and Y^n by identifying every element $\langle (x_1, y_1), \ldots, (x_n, y_n) \rangle \in Q^n$ with an edge $\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle$. Let $k = |Q|$. A graph $\{\bar{e}_1, \ldots, \bar{e}_k\} \subseteq Q^n$ is called *a full Q-graph* if it is isomorphic to graph Q and for some $l \leq n$, l-th components of n-vectors $\bar{e}_1, \ldots, \bar{e}_k$ are pairwise distinct, that is, they make up the whole set Q. Let $W \subseteq Q^n$ be the largest graph which contains no full Q-subgraph. By $E_Q(n)$ we denote its density $|W|/k^n$.

The common starting-point in [2, 11, 5, 1, 14] and in our proof of Theorem 3.1 is the bound $\omega(G^n) \leq E_Q(n)$ for any nontrivial G. So, one can try to establish the PRC by using this inequality and further estimating $E_Q(n)$. We call this approach *the extremal graph method.* The extremal graph method yields upper bounds on $\omega(G^n)$ that do not depend on the size of reply sets S and T. Moreover, the method applies for a game $G = \langle Q, S, T, \phi \rangle$ where the sets S and T are allowed to be infinite. We will call such games *unbounded.* The PRC can be generalized to unbounded games. In Section 4 it is shown that if we consider *the generalized PRC* then the extremal graph method cannot be improved. More precisely, in addition to the estimate $\omega(G^n) \leq E_Q(n)$, we have the following assertion.

For any connected graph Q there is an unbounded game $G = \langle Q, S, T, \phi \rangle$ with countable S and T such that $\omega(G^n) = E_Q(n)$.

Note that it suffices to prove or disprove the PRC for the case of connected Q.

Thus, the extremal graph method is optimal among the methods for proving the PRC that disregard the size of S and T. All the methods mentioned above fall in this class.

Remark. Simultaneously and independently of this work, there appeared the paper of Ran Raz [13] where the parallel repetition conjecture was proved in the general case. The method of [13] essentially uses the finiteness of S and T. So, there remains open a very natural combinatorial problem whether the generalized PRC is true (see discussions in Section 5).

2 Formal framework

We say that $G = \langle Q \subseteq X \times Y, S, T, \phi \rangle$ is an *(unbounded) game* if X, Y, S and T are sets, X and Y are finite, and

$$\phi : X \times Y \times S \times T \longrightarrow \{0,1\}$$

is a predicate. We regard arbitrary functions $f : X \longrightarrow S$ and $h : Y \longrightarrow T$ as *strategies* (of Players *1* and *2*). We define a *value* of the game G by

$$\omega(G) = \max_{f,h} \mathbf{P}\left[\phi(x, y, f(x), h(y)) = 1\right],$$

where the probability is taken over all randomly and uniformly chosen pairs $(x, y) \in Q$. G is *nontrivial* if $\omega(G) \neq 1$.

Given G, we define the *product game* G^n to be the game $\langle Q^n, S^n, T^n, \phi^n \rangle$ where

$$\phi^n(\bar{x}, \bar{y}, \bar{s}, \bar{t}) = \bigwedge_{i=1}^{n} \phi(x_i, y_i, s_i, t_i).$$

Here $\bar{v}$ is an n-vector $(v_1, \ldots, v_n)$. More accurately, the Cartesian power Q^n is thought of as a subset of $X^n \times Y^n$ by identifying a collection $(x_1, y_1), \ldots, (x_n, y_n)$ with a pair $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$. For players' strategies $F : X^n \longrightarrow S^n$ and $H : Y^n \longrightarrow T^n$ we designate $F(\bar{x}) = (f_1(\bar{x}), \ldots, f_n(\bar{x}))$ and $H(\bar{y}) = (h_1(\bar{y}), \ldots, h_n(\bar{y}))$.

We sometimes will consider the set Q^n, $n \geq 1$, as a bipartite graph with vertex classes X^n and Y^n. For the edge $\bar{e} \in Q^n$, we denote its vertex in X^n

by $x(\bar{e})$ and another vertex by $y(\bar{e})$. Given n-vector $\bar{v}$, we denote its i-th component by $\bar{v}|_i$.

Let $k = |Q|$. A graph $\{\bar{e}_1, \ldots, \bar{e}_k\} \subseteq Q^n$ is called *a full Q-graph* if it is isomorphic to graph Q and for some $l \leq n$, we have $\{\bar{e}_1|_l, \ldots, \bar{e}_k|_l\} = Q$. Let $W \subseteq Q^n$ be the largest graph that contains no full Q-subgraph. By $E_Q(n)$ we denote its density $|W|/k^n$.

The following proposition was first used in [2] in the case of complete Q.

Theorem 2.1 *For any nontrivial unbounded game $G = \langle Q, S, T, \phi \rangle$*

$$\omega(G^n) \leq E_Q(n).$$

Proof: Fix the strategies F and H optimal in G^n. Define $W \subseteq Q^n$ to be the set of successes of F and H in G^n, that is,

$$\begin{aligned} W \;=\; & \{(z_1, \ldots, z_n) \in Q^n : \\ & \bigwedge_{i=1}^{n} \phi\left(x(z_i), y(z_i), f_i(x(z_1), \ldots, x(z_n)), h_i(y(z_1), \ldots, y(z_n))\right) = 1\}. \end{aligned}$$

So, $\omega(G^n) = |W|/k^n$.

It suffices to show that W does not contain any full Q-subgraph. Suppose, to the contrary, that there is a full Q-graph $U = \{\bar{e}_1, \ldots, \bar{e}_k\} \subseteq W$ with l-projection full (i.e. $\{\bar{e}_1|_l, \ldots, \bar{e}_k|_l\} = Q$). Our goal is to prove that in this case G should be trivial.

Let a map I take the vertices (edges) of Q to the vertices (edges) of U. We suppose that I is an isomorphism between the graphs Q and U. Moreover, we can suppose that for any $e \in Q$ it holds $I(e)|_l = e$ and $I(X) \subseteq X^n$. Define the strategies f and h in the game G by $f(x) = f_l(I(x))$ for $x \in X$ and $h(y) = h_l(I(y))$ for $y \in Y$. The strategies f and h win always. Indeed, consider an arbitrary pair $e = (x, y) \in Q$. Let $I(e) = \bar{e}$. Then

$$\phi\left(x, y, f(x), h(y)\right) = \phi\left(x(\bar{e}|_l), y(\bar{e}|_l), f_l(x(\bar{e})), h_l(y(\bar{e}))\right) = 1$$

because $\bar{e} \in W$ and F,H win, in particular, in l-th game. ∎

The only upper bound on $E_Q(n)$ known in the general case was pointed out in [14].

Let $A = \{a_1, \ldots, a_k\}$ be a finite set and z be a variable that can be replaced with any element of A. Let $u(z)$ be an n-vector from $(A \cup \{z\})^n$ with at least one component z. Then the set

$$L = \{u(a_1), \ldots, u(a_k)\}$$

is called a *combinatorial line* in A^n. Denote by $r_k(n)$ the maximal possible density $|W|/k^n$ of a set $W \subseteq A^n$ without combinatorial lines.

Theorem 2.2 *For any bipartite graph $Q \subseteq X \times Y$, $|Q| = k$, we have*

$$E_Q(n) \leq r_k(n).$$

Proof: As easily seen, any combinatorial line in Q^n, $Q \subseteq X \times Y$, is a full Q-graph. ∎

Given a function $\rho(n)$, its amortized value is defined by

$$\bar{\rho} = \limsup_{n \to \infty} (\rho(n))^{1/n}.$$

The bound given by Theorem 2.2 is nonconstructive; for any k, $r_k(n) = o(1)$ is all what is known [9]. This fact does not allow one to prove the PRC since $\bar{r}_k = 1$.

The effective bound on $E_Q(n)$ was known only in the case if Q is a complete graph.

Theorem 2.3 ([2, 1, 5]) *If Q is a complete bipartite graph with vertex classes X and Y, i.e. $Q = X \times Y$, then*

$$\bar{E}_Q \leq \exp\left(-\frac{1}{2|X| \cdot |Y|(2 + \ln |X|)}\right).$$

3 Proof of the PRC for trees

Theorem 2.3 establishes the PRC for the class of games with Q being complete bipartite graph. In this section we prove the PRC for the case when Q is a tree.

Theorem 3.1 *If Q, $|Q| = k$, is a tree then*

$$\bar{E}_Q \leq \exp\left(-\Omega\left(\frac{1}{k^8}\right)\right). \tag{1}$$

Before the proof we state two useful propositions.

Proposition 3.1 (Chernoff's bound [4]) *Let $X_1, X_2, \ldots, X_n$ be independent random variables taking the values in $\{0,1\}$, and let $\mathbf{P}[X_i = 1] = p$ for all i. Then for any $\epsilon \in (0, p(1-p)]$ we have*

$$\mathbf{P}\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - p\right| \geq \epsilon\right] \leq 2\exp\left(-\frac{\epsilon^2}{2p(1-p)}n\right).$$

The next proposition is a quite rough but sufficient for our purposes statement of the Katona theorem.

Proposition 3.2 (Katona's theorem [10]) *Let $\mathcal{F}$ be a family of subsets of an N-element set. Suppose that any two members of $\mathcal{F}$ intersect in at least T elements. Then*

$$|\mathcal{F}| \leq N\binom{N}{(N+T)/2}.$$

Proof of Theorem 3.1:

First we give our proof in outline, but before we need some preliminaries. Arrange the edges $a_1, \ldots, a_k$ of the graph Q so that for any $j \leq k$ there is $m(j) < j$ such that a_j and $a_{m(j)}$ are adjacent. We fix a function $m(j)$ with this property. So, the graph $Q_j = \{a_1, \ldots, a_j\}$ is connected. Notice that $Q_k = Q$.

Given $\epsilon > 0$, define a subset $A_{n,\epsilon}$ of Q^n by

> $A_{n,\epsilon} = \{\bar{a} \in Q^n$: for any $S \subseteq Q$, the number of occurrences of the elements from S in $\bar{a}$ lies between $(|S|/k - \epsilon)n$ and $(|S|/k + \epsilon)n\}$.

We now give a high level overview of our proof. We have to show that every subgraph W of the graph Q^n which does not contain any full Q-subgraph has exponentially small number of edges compared to k^n. By the Chernoff bound, the density of $Q^n - A_{n,\epsilon}$ in Q^n is exponentially small (see Lemma

3.2). Therefore, it suffices to estimate the density of $W \subseteq A_{n,\epsilon}$ without full Q-subgraphs.

Consider two arbitrary adjacent edges a and b in Q which have a common vertex in $Z \in \{X, Y\}$. Call a pair edges $\bar{a}, \bar{b} \in A_{n,\epsilon}$ *good* if they have a common vertex in Z^n and, moreover, the set of a-positions in the vector $\bar{a}$ and the set of b-positions in the vector $\bar{b}$ essentially overlap. More exactly, the number of points where $\bar{a}$ has a component a and $\bar{b}$ has a component b is at least $\alpha k/n$ where α will be specified later. Suppose V is a subset of $A_{n,\epsilon}$ without a good pair of edges. Given α, suppose ϵ is sufficiently small. Then V is of exponentially small size. After routine calculation this follows from the Katona theorem (see Lemma 3.3).

Now we can use the finite induction on j from 1 to k to show that any $W \subseteq A_{n,\epsilon}$ without full Q_j-subgraphs is of exponentially small size. In every induction step, we have a full Q_{j-1}-subgraph $\{\bar{a}_1, \ldots, \bar{a}_{j-1}\} \subseteq W$ provided W is not exponentially small. We apply the previous consideration for $a = a_{m(j)}$ and $b = a_j$ in order to find $\bar{b} = \bar{a}_j$ in W which together with $\bar{a} = \bar{a}_{m(j)}$ makes up a good pair. After attaching $\bar{a}_j$ to the full Q_{j-1}-subgraph we obtain a subgraph of W isomorphic to Q_j. This Q_j-subgraph is full due to the overlapping condition in the definition of a good pair of edges (see Lemma 3.1), completing the proof.

We now turn to a detailed proof.

Further on z will denote either element of the set $\{x, y\}$. Let $\alpha = 1 - 1/k$. Define a relation R_j, $1 < j \leq k$, on $A_{n,\epsilon}$. Given $\bar{a}, \bar{b} \in A_{n,\epsilon}$, we set $\bar{a} R_j \bar{b}$ iff

$$z(a_j) = z(a_{m(j)}) \Rightarrow z(\bar{a}) = z(\bar{b}) \tag{2}$$

and

$$\exists I \subseteq \{1, \ldots, n\} : |I| > \alpha n/k \ \& \ \forall i \in I \, (\bar{a}|_i = a_{m(j)}, \, \bar{b}|_i = a_j). \tag{3}$$

The condition (2) will be called the *adjacency condition*, (3) will be referred to as the *overlapping condition*.

Given a subset $A \subseteq Q^n$, we denote its density $|A|/k^n$ by $\mu(A)$. Suppose $V \subseteq A_{n,\epsilon}$ and there are no $\bar{a}, \bar{b} \in V$ such that $\bar{a} R_j \bar{b}$. Denote the maximal possible density $\mu(V)$ by $\gamma_{j,\epsilon}(n)$. Let $\gamma_\epsilon(n) = \max_j \gamma_{j,\epsilon}(n)$.

Extending the definition of a full Q-graph, we will call a graph $\{\bar{a}_1, \ldots, \bar{a}_j\} \subseteq Q^n$ a full Q_j-graph if it is isomorphic to graph Q_j and for some $l \leq n$ it is true $\{\bar{a}_1|_l, \ldots, \bar{a}_j|_l\} = Q_j$.

Lemma 3.1 *Suppose $\epsilon < k^{-3}$. For any $j \leq k$ and $W \subseteq A_{n,\epsilon}$, if $\mu(W) > (j-1)k^2\gamma_\epsilon(n)$ then W contains a full Q_j-subgraph.*

Proof: It suffices to show that W contains a sequence $\bar{a}_1, \ldots, \bar{a}_j$ such that for any $s \leq j$ it is true

$$\bar{a}_{m(s)} R_s \bar{a}_s. \tag{4}$$

Assume this is done. The adjacency condition (2) implies that Q_j and $\{\bar{a}_1, \ldots, \bar{a}_j\}$ are isomorphic. (It is essentially that Q is a tree.) Obviously, the one-edge graph Q_1 has at least $(1/k - \epsilon)n$ "full" projections $\{a_1\}$. Due to the overlapping condition (3), attaching the second edge $\bar{a}_2$ reduces the amount of "full" projections $\{a_1, a_2\}$ in at most $(1/k+\epsilon)n - \alpha n/k$. The same is true when attaching every subsequent edge $\bar{a}_s$ (a "full" projection becomes $\{a_1, \ldots, a_s\}$). Since

$$(k-1)\left((1/k+\epsilon)n - \alpha n/k\right) < (1/k-\epsilon)n,$$

we have that $\{\bar{a}_1, \ldots, \bar{a}_j\}$ is a full Q_j-subgraph.

To prove that there is $\{\bar{a}_1, \ldots, \bar{a}_j\} \subseteq W$ with (4) for any $s \leq j$, we proceed by induction on j. The case $j = 1$ is trivial. Let $j > 1$. Define a subset U of W by

$$U = \{\bar{a} \in W : \ |\{\bar{b} \in W : \bar{a} R_j \bar{b}\}| < k\}.$$

We have $\mu(U) \leq k^2\gamma_\epsilon(n)$ because we can arrange arbitrarily U and look over all the elements of U deleting every element that is in relation R_j with any previous one. After performing this procedure twice, in the direct and reverse order, we obtain the set U' of density $\mu(U') > \mu(U)/k^2$, without any $\bar{a}$ and $\bar{b}$ in relation R_j.

Therefore, $\mu(W - U) > (j-2)k^2\gamma_\epsilon(n)$. By the induction hypothesis, $W - U$ contains a sequence $\bar{a}_1, \ldots, \bar{a}_{j-1}$ with (4) for all $s \leq j-1$. Since $\bar{a}_{m(j)} \notin U$, we have $\bar{a}_{m(j)} R_j \bar{a}_j$ for some $\bar{a}_j \in W - \{\bar{a}_1, \ldots, \bar{a}_{j-1}\}$. The sequence $\bar{a}_1, \ldots, \bar{a}_{j-1}, \bar{a}_j$ is contained in W and meets the required condition (4) for all $s \leq j$. ∎

Denote $\delta_\epsilon(n) = \mu(Q^n - A_{n,\epsilon})$. A direct corollary of Lemma 3.1 is the estimate

$$E_Q(n) < \delta_\epsilon(n) + k^3\gamma_\epsilon(n).$$

For the amortized values this gives

$$\bar{E}_Q \leq \max\{\bar{\delta}_\epsilon, \bar{\gamma}_\epsilon\}. \tag{5}$$

Lemma 3.2 $\bar{\delta}_\epsilon \leq \exp(-2\epsilon^2)$.

Lemma 3.3 *Suppose* $\epsilon < \frac{1}{4}k^{-4}$. *Then*

$$\bar{\gamma}_\epsilon \leq \exp\left(-\Omega\left(\frac{1}{k^3}\right)\right). \tag{6}$$

Set $\epsilon = \frac{1}{8}k^{-4}$. The estimate (5) and Lemmas 3.2, 3.3 imply (1), completing the proof of the theorem. It remains to prove Lemmas 3.2 and 3.3.

Proof of Lemma 3.2: Obviously, $Q^n - A_{n,\epsilon} = \bigcup_{S:S\subseteq Q} A_S$ where

$$A_S = \{\bar{a} \in Q^n : \text{the frequency of occurring elements from } S \text{ in } \bar{a} \text{ is outside the limits } [(|S|/k - \epsilon)n, (|S|/k + \epsilon)n]\}.$$

We now use the Chernoff bound. Clearly,

$$\mu(A_S) = \mathbf{P}\left[\left|\sum_{i=1}^{n} \frac{X_i}{n} - p\right| \geq \epsilon\right]$$

where $p = |S|/k$ and $\mathbf{P}[X_i = 1] = p$. It follows $\mu(A_S) \leq 2\exp(-2\epsilon^2 n)$. Summing over all $S \subseteq Q$, we have $\mu(Q^n - A_{n,\epsilon}) \leq O(\exp(-2\epsilon^2 n))$. This implies the desired bound $\bar{\delta}_\epsilon \leq \exp(-2\epsilon^2)$. ∎

Proof of Lemma 3.3: We will estimate $\gamma_{j,\epsilon}(n)$, $1 < j \leq k$.

Suppose a_j and $a_{m(j)}$ have a common vertex in $Z \in \{X, Y\}$. Recall that we can consider any subset of Q^n as a bipartite graph with vertex classes X^n and Y^n. We define B to be the number of vertices of $A_{n,\epsilon}$ in Z^n. We also define C to be the maximum possible cardinality of $V \subseteq A_{n,\epsilon}$ such that all the edges of V are incident to one and the same vertex $\bar{z} \in Z^n$, and no two edges from V have the overlapping property (3).

It is easy to see that

$$\gamma_{j,\epsilon}(n) \cdot k^n \leq B \cdot C.$$

Define D to be the minimum degree of a vertex from Z^n in the graph $A_{n,\epsilon}$. Clearly, $B \leq |A_{n,\epsilon}|/D$ and

$$\gamma_{j,\epsilon}(n) \leq \frac{C}{D}. \tag{7}$$

We will estimate C from above and D from below. Introduce the following notation. Let $Z = \{z_1, \ldots, z_r\}$ where $r = |Z|$. For $\nu \leq k$, denote $Q_\nu = \{a_s \in Q : z(a_s) = z_\nu\}$ and $k_\nu = |Q_\nu|$. It is clear that $\{Q_\nu\}_{\nu \leq r}$ is a partition of Q, and $k = \sum_{\nu \leq r} k_\nu$. For $\bar{a}, \bar{b} \in Q^n$, we have

$$z(\bar{a}) = z(\bar{b}) \Leftrightarrow \forall i \leq n \exists \nu \leq r : \bar{a}|_i, \bar{b}|_i \in Q_\nu.$$

Let $z(a_j) = z(a_{m(j)}) = z_t$.

C is the number of vectors $\bar{e}$ that can be placed into the extremal set V (see the definition of C above). We can estimate

$$C \leq C_1 \cdot 2\epsilon n C_2 \cdot n C_3 \cdot C_4 \tag{8}$$

as follows. The factor

$$C_1 = \prod_{\nu \neq t} k_\nu{}^{(\frac{k_\nu}{k} + \epsilon)n}$$

majorizes the number of choices of a component from $\bigcup_{\nu \neq t} Q_\nu$ in a vector $\bar{e}$. The other factors dominate the number of choices of a component from Q_t.

We put

$$C_2 = \binom{(\frac{k_t}{k} + \epsilon)n}{(\frac{2}{k} + \epsilon\tau_t)n}, \text{ where } \tau_t = \begin{cases} 1 & \text{if} \quad k_t > 4 \text{ or } k_t = 2 \\ 1/2 & \text{if} \quad k_t = 4 \\ -1 & \text{if} \quad k_t = 3. \end{cases}$$

The factor $2\epsilon n C_2$ bounds from above the number of allocations of the components a_j, $a_{m(j)}$ in $\bar{e}$. When it is fixed, there are at most nC_3 assignments of components a_j and $a_{m(j)}$ to allotted places where

$$C_3 = \binom{(\frac{2}{k} + \epsilon)n}{\frac{3-\alpha}{2k}n}. \tag{9}$$

To show this, we interpret a_j as 1, $a_{m(j)}$ as 0, and an assignment of a_j and $a_{m(j)}$ as a set of size $N \leq (\frac{2}{k} + \epsilon)n$. In this interpretation, the prohibition of pairs of vectors with overlapping components a_j and $a_{m(j)}$ longer than $\alpha k/n$ insures that the corresponding sets have an intersection larger than $(\frac{1-\alpha}{k} - \epsilon)n$; and the Katona theorem applies.

Finally, there are at most

$$C_4 = \begin{cases} (k_t - 2)^{(\frac{k_t - 2}{k} + \epsilon)n} & \text{if} \quad k_t > 2 \\ 1 & \text{if} \quad k_t = 2 \end{cases}$$

ways to locate the remaining components from $Q_t - \{a_{m(j)}, a_j\}$.

In the similar way we bound D from below. We have

$$D \geq D_1 \cdot D_2 \cdot D_3 \cdot D_4. \tag{10}$$

The factors are as follows.

$$D_1 = \prod_{\nu \neq t} {k_\nu}^{(\frac{k_\nu}{k} - \epsilon)n},$$

$$D_2 = \binom{(\frac{k_t}{k} - \epsilon)n}{(\frac{2}{k} - \epsilon\sigma_t)n} \text{ where } \sigma_t = \begin{cases} 1 & \text{if} \quad k_t > 3 \text{ or } k_t = 2 \\ -1 & \text{if} \quad k_t = 3. \end{cases},$$

$$D_3 = 2^{(\frac{2}{k} - \epsilon)n}, \tag{11}$$

and

$$D_4 = \begin{cases} (k_t - 2)^{(\frac{k_t - 2}{k} - \epsilon)n} & \text{if} \quad k_t > 2 \\ 1 & \text{if} \quad k_t = 2 \end{cases}.$$

Simple calculations give

$$\frac{C_1 C_4}{D_1 D_4} \leq \prod_{1 \leq \nu \leq k} {k_\nu}^{2\epsilon n} \leq e^{\frac{2\epsilon k}{e} n} \quad \text{and} \quad \frac{C_2}{D_2} < k^{2\epsilon n}.$$

This along with (7), (8), (10), (9) and (11) implies

$$\gamma_{j,\epsilon}(n) < e^{\epsilon k n} 2^{(-\frac{2}{k} + \epsilon)n} \binom{(\frac{2}{k} + \epsilon)n}{\frac{3-\alpha}{2k} n}.$$

Substituting $\alpha = 1 - 1/k$ and $\epsilon = ck^{-4}$ and assuming $c \in (0, 1/4)$, we obtain the desired bound

$$\bar{\gamma}_{j,\epsilon} \leq \exp\left(-\Omega\left(\frac{1}{k^3}\right)\right)$$

for arbitrary j. Thus, (6) follows.

■

■

4 The extremal density and unbounded games

Theorem 2.1 shows that $E_Q(n)$ is an upper bound on $\omega(G^n)$ for any nontrivial unbounded game G. In fact, this bound cannot be improved. The following theorem shows that the generalized PRC for unbounded games is equivalent to the question whether or not $\bar{E}_Q(n) < 1$ for any Q.

Theorem 4.1 *For any connected bipartite graph $Q \subseteq X \times Y$ there is a game $G = \langle Q, S, T, \phi \rangle$ with S and T countable such that*

$$\omega(G^n) = E_Q(n).$$

Proof: Given a connected bipartite graph $Q \subseteq X \times Y$, we have to construct $\phi \subseteq X \times Y \times S \times T$. What are the sets S and T will be clear from our construction. Let $W_n \subseteq Q^n$ be the largest graph without full Q-subgraphs. We set

$$\phi = \{\langle \bar{x}|_i, \bar{y}|_i, (i, \bar{x}), (i, \bar{y}) \rangle : \exists i, n, \bar{e} \text{ s.t. } i \leq n, \bar{e} \in W_n, x(\bar{e}) = \bar{x}, y(\bar{e}) = \bar{y}\},$$

thereby defining an unbounded game G.

Consider the strategies F and H in G^n defined by $f_i(\bar{x}) = (i, \bar{x})$ and $h_i(\bar{y}) = (i, \bar{y})$. Obviously, F and H win on W_n, therefore, $E_Q(n) \leq \omega(G^n)$.

To prove $\omega(G^n) \leq E_Q(n)$, by Theorem 2.1 it suffices to prove that G is nontrivial. Suppose, to the contrary, that some strategies $f : X \longrightarrow S$ and $h : Y \longrightarrow T$ always win, that is, for any $(x, y) \in Q$ we have $\langle x, y, f(x), h(y) \rangle \in \phi$. From the connectivity of Q it follows that there are n and $l \leq n$ such that for any $x \in X$ it holds $f(x) = (l, \hat{x})$ for some $\hat{x} \in X^n$ with $\hat{x}|_l = x$ and, similarly, for any $y \in Y$ it holds $h(y) = (l, \hat{y})$ for some $\hat{y} \in Y^n$ with $\hat{y}|_l = y$. Moreover, there are $\bar{e}_1, \ldots, \bar{e}_k$ in W_n such that $(x, y) \in Q$ implies $\hat{x} = x(\bar{e}_j)$ and $\hat{y} = y(\bar{e}_j)$ for some $j \leq k$. It is clear that the graph $\{\bar{e}_1, \ldots, \bar{e}_k\}$ is isomorphic to the graph Q and its l-th projection gives the whole Q. Thus, $\{\bar{e}_1, \ldots, \bar{e}_k\}$ is a full Q-subgraph of W_n yielding the contradiction. ∎

5 Open problems

1. Prove the generalized PRC for cycles (of even length). After Theorems 2.3 and 3.1, the first unexplored case is when $|X| = |Y| = 3$ and $Q = C_6$ is a cycle of length six.

2. Let $D_Q(n)$ denote the maximal possible density of $W \subseteq Q^n$ without subgraphs isomorphic to Q. Obviously, $D_Q(n) \leq E_Q(n)$. Prove, at least, $\bar{D}_Q < 1$ for any Q. This natural problem seems to be of independent interest. If Q is a complete graph, the inequality $\bar{D}_Q < 1$ is true by the Zarankievicz theorem. If Q is a tree, this follows from the well-known fact that every graph on v vertices with at least tv edges contains as a subgraph any tree of size t. The only what I know in the general case is $D_Q(n) = o(1)$.

Acknowledgments. Theorem 4.1 appeared as an answer to the remark of Alex Russell on the possibility of inverting the relation between the winning probability and a Ramsey type function. I thank to him and also to Rostyslav Hryniv and Yaroslav Vorobets for several fruitful discussions. I am grateful to Alexander Razborov for useful technical suggestions.

References

[1] N. Alon, Probabilistic methods in extremal finite set theory, to appear in *the Proc. of the Conference on Extremal Problems for Finite Sets*, (G. O. H. Katona et al. Eds., Visegrad, Hungary, 1991) 1–13.

[2] J. Cai, A. Condon, and R. J. Lipton, Playing games of incomplete information, *Theoret. Comput. Sci.* **103** (1992) 25–38.

[3] J. Cai, A. Condon, and R. J. Lipton, $PSPACE$ is provable by two provers in one round, in: *Proc. of the 6th Ann. Conference on Structure in Complexity Theory* (1991) 110–115.

[4] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Math. Stat.*, 23:493–509, 1952.

[5] U. Feige, On the success probability of the two prover in one round proof systems, in: *Proc. of the 6th Ann. Conference on Structure in Complexity Theory* (1991) 116–123.

[6] U. Feige and L. Lovász, Two-prover one-round proof systems: their power and their problems, in: *Proc. of the 24th ACM Ann. Symp. on the Theory of Computing (STOC)* (1992) 733–744.

[7] L. Fortnow, Complexity-theoretic aspects of interactive proof systems, Ph.D. thesis, Tech. Report #MIT/LCS/TR-447, Massachusetts Institute of Technology, 1989.

[8] L. Fortnow, J. Rompel, and M. Sipser, On the power of multi-prover interactive protocols, in: *Proc. of the 3rd Ann. Conference on Structure in Complexity Theory* (1988) 156–161, Erratum in: *Proc. of the 5th Ann. Conference on Structure in Complexity Theory* (1990) 318–319.

[9] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *Journal d'Analyse Mathematique* **57** (1991) 64–119.

[10] Gy. Katona. Intersections theorems for systems of finite sets. *Acta Math. Hungar.*, 15:329–337, 1964.

[11] D. Lapidot and A. Shamir, A one-round, two-prover, zero-knowledge protocol for NP, in: *CRYPTO'91.*

[12] D. Peleg, On the maximal number of ones in zero-one matrices with no forbidden rectangles, manuscript, 1990.

[13] R. Raz, A parallel repetition theorem, manuscript, 1994.

[14] O. Verbitsky, Towards the parallel repetition conjecture, in: *Proc. of the 9th Ann. Conference on Structure in Complexity Theory* (1994) 304–307.