Monotone Circuits for Connectivity Require Super-logarithmic Depth

Mauricio Karchmer
Avi Wigderson*

Inst. of Mathematics and Computer Science
Hebrew University
Jerusalem, Israel 91904

Abstract

We prove that every monotone circuit which tests \(s-t\)-connectivity of an undirected graph on \(n\) nodes has depth \(\Omega(\log^2 n)\). This implies a superpolynomial \((n^{\Omega(\log n)})\) lower bound on the size of any monotone formula for \(s-t\)-connectivity.

The proof draws intuition from a new characterization of circuit depth in terms of communication complexity. It uses counting arguments and Extremal Set Theory.

Within the same framework, we also give a very simple and intuitive proof of a depth analogue of a theorem of Krapchenko concerning formula size lower bounds.

1 Introduction

The circuit complexity of Boolean functions has been studied for 40 years, but its main problem remains unsolved: we have no example of a simple function (say in NP) that requires super linear circuit size or super logarithmic (bounded fan-in) circuit depth. The reason is, perhaps, that although the circuit model is elegantly simple, our understanding of the way it computes is, at the most, vague.

In the last years, however, advance has been surprisingly fast. On the one hand, results of

*part of this work was done while the authors were visiting UC Berkeley supported by an NSF grant DCR-8612563. The second author also wishes to acknowledge support from the Alon Fellowship.
Andreev [An] and Razborov [Ra], improved by
Alon and Boppana [AB], give exponential size
lower bounds for monotone circuits. On the
other hand, results of, by now a long list of au-
thors (e.g. [A],[FSS],[Y1],[H]), give expo-
ential size lower bounds for constant depth cir-
cuits. More than the results themselves, per-
haps the main contribution of the mentioned
papers has been the development of some gen-
eral techniques for proving lower bounds, such
as random restrictions and circuit approxima-
tion. These techniques, however, turn out to
be hard to apply to other problems so that
new ideas have been sought.

In this paper we show the equivalence be-
tween circuit depth and the communication
complexity of a certain related problem 1. We
believe that the later model is much more ap-
pealing for both showing and understanding
upper bounds, as well as for proving lower
bounds. This characterization is reminiscent
of, but somehow more explicit and intuitive
than, the well known relationship between cir-
cuits and alternating machines [Ru]. This
characterization allows us to view compu-
tation top-down (from output to input) and
apply such techniques as random restrictions
in that direction (rather than the common
bottom-up approach). We argue the rele-
ance of this model by presenting a very sim-
ple proof of a depth analogue of a theorem of
Krapchenko, and by proving the first super-
logarithmic (in the size of the circuit) depth
lower bound for monotone circuits.

Though the mentioned results of Andreev
and Razborov give exponential (in \( \log n \))
depth lower bounds for monotone circuits
computing certain functions, the depth lower
bound is always logarithmic in the size bound.

That is, the techniques apply to size rather
than to depth. We present a technique which
captures, in a strong way, the essence of circuit
depth. We give here a tight \( \Omega(\log^2 n) \) depth
bound for \( st \)-connectivity 2, a function which
has \( O(n^3 \log n) \) size, \( O(\log^2 n) \) depth mon-
tone circuits. As a consequence, we get non-
polynomial \( (n^{\Omega(\log n)}) \) size lower bounds for
monotone formulas computing \( st \)-connectivity
and hence separate the monotone analogues of
\( NC^1 \) and \( AC^1 \).

While our proof bears no obvious similarity
to the methods of Razborov and Andreev,
we point out the important role that (dif-
terent nontrivial results from extremal set the-
ory play in both cases.

It is interesting to note here the different
character of the connectivity and majority
functions in the Boolean and arithmetic mono-
tone circuits models. Shamir and Snir [ShS]
showed an \( \Omega(\log^2 n) \) depth bound for both
functions in the arithmetic model. The dif-
culty in applying these techniques to Boolean
circuits are the axioms \( x \lor xy = x \) and its dual,
which do not hold in fields. Indeed, Valiant
[V] (by probabilistic methods) and [AKS] (by
explicit constructions) showed that these ax-
ioms make a difference for the majority func-
tion which admits \( O(\log n) \) depth monotone
Boolean circuits. Our result says that, unlike
for majority, for connectivity the situation in
the Boolean case is very similar to the arith-
metic one.

It is worthwhile to mention that our results
apply to undirected graph \( st \)-connectivity,
a function that, in some models, is easier

\(^1\) We were told that Yannakakis indepen-
dently discovered this equivalence which is im-
licit in [KPPY].

\(^2\) We present here and improved and simpli-
fied version of an early result of ours giving a
\( \Omega(\log^2 n / \log \log n) \) bound. This was possible
after J. Hastad formulated and proved lemma
4.1. A similar improvement was independently
discovered by R. Boppana.
than its directed version. For example, see [AKLLR] for some relevant evidence. More recently, Ajtai and Fagin [AF] show that, while undirected st-connectivity is definable in monadic second order logic, the directed case is not.

The paper is organized as follows: In §2 we define the communication game and show its equivalence to circuit depth; in §3 we give a simple proof of a theorem of Krapchenko. In §4 we give the lower bound for connectivity.

2 Communication Complexity and Circuit Depth

In this section we show the equivalence between circuit depth and a problem in communication complexity. We will be considering circuits over the basis \( \{\lor,\land,\neg\} \) where \( \{\lor,\land\} \)-gates have fanin 2 and \( \neg \)-gates are only applied to input variables. For a function \( f \), \( d(f) \) is the minimum depth of a circuit computing \( f \).

Let \( B_0, B_1 \subseteq \{0,1\}^n \) such that \( B_0 \cap B_1 = \emptyset \). Consider the following game between players I and II: Player I gets \( x \in B_1 \) while player II gets \( y \in B_0 \); their goal is to find a coordinate \( i \) such that \( x_i \neq y_i \). Let \( C(B_1, B_0) \) be the minimum number of bits they have to communicate in order for both to agree on such coordinate. Note that unlike standard problems in communication complexity [Y], the task of the players here is to solve a search, rather than a decision, problem.

**Theorem 2.1** For every function \( f : \{0,1\}^n \to \{0,1\} \) we have

\[
d(f) = C(f^{-1}(1), f^{-1}(0))
\]

**Proof:** Follows from the following two lemmas.

**Lemma 2.1** For all functions \( f \), and all \( B_0, B_1 \subseteq \{0,1\}^n \) such that \( B_0 \subseteq f^{-1}(0) \) and \( B_1 \subseteq f^{-1}(1) \) we have

\[
C(B_1, B_0) \leq d(f)
\]

**Proof:** By induction on \( d(f) \).

If \( d(f) = 0 \) then \( f \) is either \( x_i \) or \( \bar{x}_i \). In either case, we have that for all \( x \in B_1 \) and \( y \in B_0 \), \( x_i \neq y_i \) so that \( i \) is always an answer and \( C(B_1, B_0) = 0 \).

For the induction step we suppose that \( f = f_1 \land f_2 \) (the case \( f = f_1 \lor f_2 \) is treated similarly) so that \( d(f) = \max(d(f_1), d(f_2)) + 1 \). Let \( B_j^0 = B_0 \cap f_j^{-1}(0) \) for \( j = 1, 2 \). By induction we have that \( C(B_1, B_j^0) \leq d(f_j) \) for \( j = 1, 2 \). Consider the following protocol for \( B_1 \) and \( B_0 \): II sends a 0 if \( y \in B_j^0 \), otherwise he sends a 1; the players then follow the best protocol for each of the subcases. We have

\[
C(B_1, B_0) \leq 1 + \max_{j=1,2}(C(B_1, B_j^0))
\]

\[
\leq 1 + \max_{j=1,2}(d(f_j))
\]

\[
= d(f)
\]

The converse is as follows:

**Lemma 2.2** Let \( B_0, B_1 \subseteq \{0,1\}^n \) such that \( B_0 \cap B_1 = \emptyset \). Then, there exists a function \( f \) with \( B_0 \subseteq f^{-1}(0) \) and \( B_1 \subseteq f^{-1}(1) \) such that

\[
d(f) \leq C(B_1, B_0)
\]

541
Proof: By induction on \( C(B_1, B_0) \).

If \( C(B_1, B_0) = 0 \) then there exists an \( i \) such that \( \forall x \in B_1 \) and \( \forall y \in B_0, x_i \neq y_i \). It is clear that \( \forall x', x'' \in B_1 \) we have \( x'_i = x''_i \) and the same holds for all \( y', y'' \in B_0 \). Without loss of generality \( x_i = 1 \) so that letting \( f = x_i \) we have \( B_0 \subseteq f^{-1}(0) \) and \( B_1 \subseteq f^{-1}(1) \).

To prove the induction step, we assume that \( B_2 \) sends the first bit (the other case is treated similarly). For some partition \( B_0 = B_0^1 \cup B_0^2 \), \( B_1 \) sends a 0 if \( y \in B_0^1 \), a 1 otherwise; the players then continue with the best protocol for each of the subcases and

\[
C(B_1, B_0) = 1 + \max_{j=1,2}(C(B_1, B_0^j))
\]

By induction, there exist \( f_1, f_2 \) such that \( B_0^j \subseteq f_j^{-1}(0) \), \( B_1 \subseteq f_j^{-1}(1) \) and \( d(f_j) \leq C(B_1, B_0^j) \) for \( j = 1, 2 \). Taking now \( f = f_1 \land f_2 \) we have

\[
B_0 \subseteq f_1^{-1}(1) \cap f_2^{-1}(1) = f^{-1}(1)
\]

\[
B_0 = B_0^1 \cup B_0^2 \subseteq f_1^{-1}(0) \cup f_2^{-1}(0) = f^{-1}(0)
\]

and

\[
d(f) \leq 1 + \max_{j=1,2}(d(f_j)) \leq 1 + \max_{j=1,2}(C(B_1, B_0^j)) = C(B_1, B_0)
\]

Theorem 2.2 For every monotone function \( f \) we have

\[
d_m(f) = C_m(\min(f), \max(f))
\]

Proof: Note that in the base case of lemma 2.1, if the circuit is monotone, we always find a coordinate \( i \) such that \( x_i = 1 \) while \( y_i = 0 \). In the other hand, if the protocol always gives a coordinate \( i \) with the above property, lemma 2.2 gives a monotone circuit.

Let \( x \in f^{-1}(1) \) be the characteristic vector of a subset \( p \subseteq [n] \). Similarly, let \( y \in f^{-1}(0) \) be the characteristic vector of the complement of a subset \( q \subseteq [n] \) (i.e. \( i \in q \iff y_i = 0 \)). By the above argument, it is clear that the answer of the protocol will be an element of \( p \land q \). The theorem follows by noticing that it is enough to give a protocol for \( (\min(f), \max(f)) \) because the players, in case they get inputs \( p' \) and \( q' \), can always behave as if they got \( p \subseteq p' \) and \( q \subseteq q' \) where \( p \in \min(f) \) and \( q \in \max(f) \).

For proving lower bounds for the communication game, it may be convenient to have

\[3[n] = \{1, \ldots, n\}\]

542
more structure in the way the players behave. We would like to synchronize the protocol so that the players communicate in rounds where players I and II send messages of length at most $a_I, a_{II}$ respectively:

**Theorem 2.3** Let $D$ be a protocol in rounds where at each round player I sends $a$ bits and player II responds with $2^a$ bits ($a \leq \log n$). Then for any function $f$, the number $k$ of rounds satisfies

$$k \leq \frac{d(f)}{a}$$

for the general game and

$$k \leq \frac{d_m(f)}{a}$$

for the monotone one.

**Proof:** Let $C$ be the best circuit for $f$. The idea is to simulate $a$ layers of $C$ with a round of $D$. Divide $C$ into stages of depth $a$ each and look at the subcircuits of each stage. Each one computes a function which depends on at most $2^a$ wires and, thus, can be represented in CNF form with less than $2^{2^a}$ clauses, each of length $2^a$. Following the proof of lemma 2.1, it is easy to see that such CNF representation can be simulated by a round where player II sends $2^a$ bits and player I sends a bits. The same holds for the monotone case.

Of course, in theorem 2.3, the roles of I and II can be switched so that I and II send $2^a$, a bits per round respectively.

### 3 Krapchenko’s bound

As a nice application of theorem 2.1, we give a simple proof of a depth analogue of a theorem of Krapchenko. Let $C_n$ be the graph of the $n$-cube with vertex set $\{0, 1\}^n$ and two nodes adjacent iff they differ in one coordinate. Any subset $A$ of edges induce a graph $G_A$ of $C_n$ in the natural way. For a graph $G_A$ and a node $x$, we denote $d_A(x), N_A(x)$ as the degree of $x$ in $A$ and the set of neighbors of $x$ in $A$ respectively. We drop the subindex $A$ if no confusion arises. Let $E$ denote expectation with uniform distribution.

**Theorem 3.1 (Krapchenko)** Let $B_0, B_1 \subseteq \{0, 1\}^n$ such that $B_0 \cap B_1 = \emptyset$. Let $A = C_n \cap (B_0 \times B_1)$. Then, for every function $f$ with $B_0 \subseteq f^{-1}(0)$ and $B_1 \subseteq f^{-1}(1)$ we have

$$d(f) \geq \log \frac{|A|^2}{|B_0||B_1|}$$

**Proof:** Fix a protocol $D$ for the communication game and let $C(x,y)$ be the number of bits $D$ uses on inputs $x,y$. We will prove that for $(x,y)$ taken uniformly from $A$ we have

$$E(C(x,y)) \geq \log \frac{|A|^2}{|B_0||B_1|} \quad (\ast)$$

By lemma 2.1, we get $d(f) \geq E(C(x,y))$. We view $\ast$ as follows: Write

$$\log \frac{|A|^2}{|B_0||B_1|} = \log \frac{|A|}{|B_0|} + \log \frac{|A|}{|B_1|}$$

and notice that $|A|/|B_0|$ and $|A|/|B_1|$ are the average degree of nodes in $B_0$ and $B_1$ respectively. In what follows, we will claim that the number of bits player I sends is at least the logarithm of the average degree of nodes in $B_1$ (similarly with player II). Intuitively, this is so because even if player I knows $y$, he needs $\log d(y)$ bits to tell II which $x$ he has.

We now proceed formally. For $(x,y) \in A$, let $b_I(x,y), b_{II}(x,y)$ be the number of bits players I, II send when the input to the protocol
is \((x, y)\). We have

\[
E(C(x, y)) = \frac{1}{|A|} \left[ \sum_{(x, y) \in A} (b_1(x, y) + b_2(x, y)) \right]
\]

\[
= \frac{1}{|A|} \left[ \sum_{x \in B_1} \sum_{y \in N(x)} b_2(x, y) + \sum_{y \in B_0} \sum_{x \in N(y)} b_1(x, y) \right]
\]

We claim:

- For any \(x \in B_1\), \(\sum_{y \in N(x)} b_2(x, y) \geq d(x) \log d(x)\). This is so because, even if \(\Pi\) knows \(x\), he has to tell \(I\) which \(y\) he has.

- Similarly, \(\forall y \in B_0\) we have \(\sum_{x \in N(y)} b_1(x, y) \geq d(y) \log d(y)\).

We now conclude,

\[
E(C(x, y)) \geq \frac{1}{|A|} \left( \sum_{x \in B_1} d(x) \log d(x) \right) + \frac{1}{|A|} \left( \sum_{y \in B_0} d(y) \log d(y) \right)
\]

\[
\geq \frac{1}{|A|} \left( \sum_{x \in B_1} \frac{|A|}{|B_1|} \log \frac{|A|}{|B_1|} + \sum_{y \in B_0} \frac{|A|}{|B_0|} \log \frac{|A|}{|B_0|} \right)
\]

\[
= \log \frac{|A|^2}{|B_1||B_0|}
\]

where the last inequality follows from the convexity of \(x \log x\).

\[\square\]

4 A Lower Bound for Connectivity

In this section we give a \(\Omega(\log^2 n)\) depth lower bound for monotone circuits computing undirected graph \(st\)-connectivity. This section is organized as follows: In §4.1 we give some intuition and we state the main theorem; In §4.2 we give some definitions and useful lemmas; finally, in §4.3 we give the proof of the theorem.

4.1 Intuition

The function \(st\)-connectivity receives as input an undirected graph and two distinguished vertices \(s\) and \(t\), and tests whether there is a path from \(s\) to \(t\) or not. The function is obviously monotone with minterms corresponding to minimal \(st\)-paths, and maxterms corresponding to minimal \(st\)-cuts. We view \(st\)-paths as ordered sets of vertices excluding \(s\) and \(t\). We view \(st\)-cuts as a partition of the set of vertices into two subsets, one containing \(s\) and the other containing \(t\). The minimal cut contains all edges between the two subsets. This is the same as having a coloring \(q : V \to \{0, 1\}\) where \(q(s) = 0\) and \(q(t) = 1\). The game thus is as follows: Player I gets an \(st\)-path while player II gets a coloring of the nodes. Their goal is to find a bichromatic edge in the path.

Let us look at the protocol based on the idea of raising the adjacency matrix of the graph to the \(n^\text{th}\) power: Player I sends the name of the middle vertex on his path, player II responds with the color of that vertex. The players then continue recursively on the half path where a bichromatic edge is ensured to exist. Note that the protocol requires \(O(\log n)\) rounds in each of which I sends \(\log n\) bits and II sends just one.

The crucial observation is that, even if player II would be allowed to send \(O(n^4)\) bits each round (instead of one bit as in the protocol), the players will still need many rounds. Basically, this is because II doesn't know much about the nodes in I's path. If he sends \(O(n^4)\) bits and the path is of length \(O(n^4)\) then the
probability that I get valuable information from II is negligible. If we could prove a \( \Omega(\log n) \) lower bound for the number of rounds needed, we will be able to use theorem 2.3 to get the promised \( \Omega(\log^2 n) \) depth lower bound for circuits.

Note the asymmetry between players I and II. Indeed, if the roles of both players were switched so that player I would be the one which sends \( O(n^\epsilon) \) bits per round, they would be able to solve the problem in a constant number of rounds. This is consistent with the intuition gotten by Shamir and Snir in [ShS].

Define \( s_{\text{conn}}(l) \) as the restriction of \( s \)-connectivity to the case where player I gets a path of length \( l \). We state the main theorem of this section:

**Theorem 4.1** Let \( l \leq n^{1/10} \). There exists an \( 0 < \epsilon < 1/2 \) such that if \( D \) is a \( k \)-round protocol for \( s_{\text{conn}}(l) \) where at each round, player I sends \( c \log n \) bits and player II sends \( n^\epsilon \) bits, then \( k \geq \log l \).

**Corollary 4.1** Any monotone circuit computing \( s \)-connectivity requires \( \Omega(\log^3 n) \) depth.

**Proof:** Follows from theorems 2.3 and 4.1 by taking \( l = n^{1/10} \). \( \square \)

**Open 4.1** Prove a similar lower bound for Connectivity.

**Corollary 4.2** Any monotone formula for \( s_{\text{conn}}(l) \) has size \( n^{\Omega(\log n)} \).

**Proof:** Follows by noting that the relation \( d(f) = O(\log L(f)) \) holds also in the monotone case. \( \square \)

To prove theorem 4.1, we will assume, for contradiction, the existence of a \( k \)-round protocol \( (k < \log l) \) good for a large family of all possible paths and a large family of all possible colorings. We will pick a large subset of the paths and colorings for which players I and II sent the same message in the first round. We will give some extra information (by applying a random restriction to the coloring of the nodes) to both players so as to get smaller, yet nicer, subsets which are in O(1)-correspondence with a family of paths shorter in length (but of higher quality) and a family of colorings of fewer nodes. The fact that the original protocol had \( (k - 1) \) rounds to go, will allow us to find a \( (k - 1) \)-round protocol for the smaller families. Repeating this \( \Omega(\log n) \) times will give us a protocol without communication that solves a problem which cannot be solved without any messages.

Note the top-down structure of the proof; essentially, the argument shows that, if the output of a circuit is in some sense complex and, as long as we do not go too far down the circuit, there is a wire which computes a complex subfunction.

### 4.2 Notation and Definitions

Denote by \( \Pi^n \subseteq [n]^l \) the set of all paths on \( [n] \) of length \( l \). \( |\Pi^n| = (n)_l \) where

\[
(n)_l = n(n-1)\ldots(n-l+1)
\]

An interval \( I \subseteq [l] \) is a subset of consecutive integers. For a path \( p \in \Pi^n \) and an interval \( I \subseteq [l] \), \( p_{\mid I} \) is the subpath of \( p \) in the interval \( I \). For \( P \subseteq \Pi^n \), \( P_{\mid I} = \{ p_{\mid I} : p \in P \} \) is the projection of \( P \) into \( I \). Note that \( P_{\mid I} \subseteq \Pi^n_{\mid I} \). Conversely, for \( p \in \Pi^n_{\mid I} \), \( p \subseteq \Pi^n \) and an interval \( I \), let \( \text{Ext}_{p_{\mid I}}(p) = \{ \bar{p} \in P : \bar{p}_{\mid I} = p \} \) be the set
of extensions of $p$ in $P$ within $I$. We will drop
the subindices $P$ and $I$ if no confusion arises.
For $p \in \Pi^n$, the support of $p$, supp($p$), is
defined as the set of nodes contained in $p$. When
no confusion arises, we will denote supp($p$) by $p$.
For $P \subseteq \Pi^n$, supp($P$) = \{supp($p$) : $p \in P$\}.
Given a partition of $[I]$ into two intervals $L$ and $R$, we will denote a path $p \in \Pi^n$ by $(p_L, p_R)$
where each entry is the projection of $p$ into the
respective interval.

Similarly, for a coloring $q \in \{0, 1\}^n$ and a
subset $T \subseteq [n]$, $q_T$ is the projection of $q$ into $T$
and for $Q \subseteq \{0, 1\}^n$, $Q_T = \{q_T : q \in Q\}$ is
the projection of $Q$ into $T$. For $q \in \{0, 1\}^{|T|}$, $Q \subseteq
\{0, 1\}^n$ and a subset $T \subseteq [n]$, let $Ext_{Q,T}(q) =
\{q_T \in Q : q_T = q\}$ be the set of extensions of $q$
in $Q$ within $T$ (again, we drop the subindices $Q$ and $T$
whenever possible). For a restriction $\rho : [n] \rightarrow \{0, 1, *, \}$, we will denote by $Q_\rho$ the
set of colorings in $Q$ consistent with $\rho$, (i.e.
$q \in Q : \rho(i) \neq * \Rightarrow \rho(i) = q_i$).

For a subset $A$ of a universe $\Omega$, the density
of $A$, $\mu(A)$, is defined as $|A|/|\Omega|$. In what follows,
we will work with densities rather than with cardinalities.

We will need the following combinatorial lemma due to J. Hastad: Let $H \subseteq A_1 \times \ldots \times A_k$ and for $v \in A_i$, let $Ext_{A_i}(v) = \{u \in H : u_i = v\}$. Note that, though $Ext_{A_i}(v) \subseteq H$, $Ext_{A_i}(v)$ may be considered as a subset of $H/A_i = A_1 \times \ldots \times A_{i-1} \times A_{i+1} \times \ldots \times A_k$. In what follows, we will consider $Ext_{A_i}(v)$ as
a subset of $H/A_i$.

**Lemma 4.1** Let $H \subseteq A_1 \times \ldots \times A_k$. Let $B_i =
\{u \in A_i : \mu(Ext_{A_i}(u)) \geq \mu(H)/2k\}$. Then

$$\prod_{i=1}^{k} \mu(B_i) \geq \frac{\mu(H)}{2}$$

**Proof:** Say that a member $(u_1, \ldots, u_k)$ of $H$
is bad if for some $i$, $u_i \notin B_i$. Let $H$ be the set
of bad elements in $H$. We have

$$\mu(H) \leq \sum_{i=1}^{k} \mu(\cup_{u \in B_i} Ext_{A_i}(u))$$

$$< \sum_{i=1}^{k} \frac{\mu(H)}{2k} = \frac{\mu(H)}{2}$$

the lemma follows immediately, by noting that

$$\prod_{i=1}^{k} \mu(B_i) \geq \mu(H).$$

**Corollary 4.3** If $k = 2$, then $\exists i$ such that

$$\mu(B_i) \geq \left( \frac{\mu(H)}{2} \right)^{1/2}.$$

**Corollary 4.4** $Pr(\mu(B_i) < \left( \frac{\mu(H)}{2} \right)^{2/k}) <
\frac{1}{2}$ for $i$ chosen randomly from $\{1, \ldots, k\}$.

We will also need the following result in
extremal graph theory which is based on the
Kruskal-Katona theorem. Let $F \subseteq 2^{[n]}$. $F$ is
called an ideal if $A \in F$ and $B \subseteq A$ implies
that $B \in F$. For $0 \leq s \leq n$, $P_s(F)$ is the
probability that a random $s$-subset $S$ is in $F$.
We present the following lemma of Bollobas
and Thomason:

**Lemma 4.2** (BTh) Let $F \subseteq 2^{[n]}$ be an ideal.
Then for $0 \leq s < r \leq n$ we have

$$P_s(F)^s \geq P_r(F)^r$$

Intuitively, the lemma says that the probabil-
ability of the different slices of an ideal
decreases exponentially fast with their distance.
Our main use of the lemma is the following
corollary:

546
Corollary 4.5 Let $S \subseteq [n]^{(s)}$ be a family of $s$-subsets of $[n]$ with $\mu(S) \geq \alpha$. Take a random $t$-subset $T$ of $[n]$. The probability that $T$ does not contain any element of $S$ is at most $(1 - \alpha)^{t/s}$.

Proof: Say $T$ is good if it contains an element of $S$, $T$ is bad otherwise. It is easy to see that the ideal $I$ generated by the bad $T$'s misses all of $S$ so that $P_t(I) \leq (1 - \alpha)$. By lemma 4.3 we get

$$\Pr(T \text{ is bad}) = P_t(I) \leq (1 - \alpha)^{t/s}$$

It is clear that the two claims imply $-H(0, t_{max})$ which in turn implies our theorem.

4.3 The proof

Proof: [of theorem 4.1]

In what follows, all our protocols will be synchronized so that at each round $I$ sends $\epsilon \log n$ bits and $II$ responds with $n^t$ bits. The existence of $\epsilon$ will be clear from the proof, though one can check that $\epsilon = 1/10$ suffices. We will define a sequence of problems of different sizes as follows: We first define the parameters of the problems, let $t_{max} = \log l - 1$.

Let $V_0 = [n]$ and $V_{i+1} \subseteq V_i$ where $|V_i| = n_i$ and $n_{i+1} = n_i - 4n_i^{1/2}$. Note that

$$n/2 \leq n_t \leq n \quad \text{for} \quad t \leq t_{max}$$

Let $l_0 = l$ and $l_{i+1} = l_i/2$ and note that

$$2 \leq l_t \leq l \quad \text{for} \quad t \leq t_{max}$$

Consider the following property:

$H(t, k)$: There exist a collection of paths $P^t \subseteq \Pi_{t}^{n_t}$ of length $l_t$ over $V_t$, and a collection of colorings $Q^t \subseteq \{0,1\}^{n_t}$ of $V_t$, with $\mu(P^t) \geq n^{-\epsilon}$ and $\mu(Q^t) \geq 2^{-2n^{t}}$ such that there exists a $k$-round protocol $D^t$ good for $(P^t, Q^t)$.

We will prove the following two claims:

Claim 4.1 For $t \leq t_{max}$ $-H(t, 0)$.

Claim 4.2 For $t \leq t_{max}$ $H(t, k) \rightarrow H(t + 1, k - 1)$
Lemma 4.3  $P$ is either $L$-good or $R$-good.

Proof: We have $P \subseteq \Pi_{i+1}^{n_{i+1}} \times \Pi_{i+1}^{n_{i+1}}$. Note that $\mu(P) \geq n^{-2e}(1 - n^{-4/5})$ by the choice of $l$. The lemma follows using corollary 4.3 and noting that $n^{-2e}(1 - n^{-4/5})/4 \geq n^{-3e}$. $\blacksquare$

Without loss of generality, assume that $P$ is $L$-good and let $A$ be the set of paths in $P_L$ with many extensions. The next lemma is the heart of our argument:

Lemma 4.4 There exists a restriction $\rho : V_i \rightarrow \{0, 1, *\}$ with $|\rho^{-1}(*)| = n_{i+1}$ such that the following properties hold:

G1: If $Q = Q_\rho$ then $\mu(Q) \geq 2^{-2(i+1)n^*}$.

G2: $\exists \bar{P} \subseteq P$ such that

- $\forall p \in \bar{P}, p_L \subseteq \rho^{-1}(*),$ and $p_R \subseteq \rho^{-1}(1)$
- $\forall p, p' \in \bar{P}$, $p_L \neq p'_L$
- $\mu(\bar{P}) \geq n^{-e}$.

Assuming the lemma is true, we will finish the proof of the second claim:

Let $V^{i+1} = \rho^{-1}(*), Q^{i+1} = Q_{\rho^{-1}(*)}$, and $P^{i+1} = \bar{P}_L$. Note that there is a natural 1-1 correspondence between $Q^{i+1}$ and $Q$ and between $P^{i+1}$ and $\bar{P}$. Also note that $\forall q \in Q$ and $\forall p \in \bar{P}$ any bichromatic edge lies in the interval $P$. The protocol $D^{i+1}$ on $(P^{i+1}, Q^{i+1})$ simulates the protocol $D^i$ on $(\bar{P}, \bar{Q})$ by following the behaviour of the associated path and coloring. $\blacksquare$

Proof: [of lemma 4.4]

In what follows, we denote $V = V_t, v = n_t, l = l_t, l' = l_{i+1}$ for simplicity sake. The existence of a good restriction will be shown by probabilistic methods. We will pick $\rho$ uniformly from the set of all restrictions with $|\rho^{-1}(*)| = v - 4\sqrt{n}$ and $\Pr(\rho(x) = 0 | \rho(x) \neq *) = 1/2$, and show that, with positive probability, the conditions of the lemma are fulfilled. Specifically, we will show that $\Pr(\neg G1) + \Pr(\neg G2) \leq 1/2 + o(1)$.

Let us start with $G1$: Intuitively, the following lemma says that, with high probability, $\rho$ does not give player I too much information about the colors of nodes in $\rho^{-1}(*)$.

Lemma 4.5  $\Pr(\mu(Q_\rho) < 2^{-2(i+1)n^*}) \leq \frac{1}{2} + o(1)$.

Proof:

Let $\alpha = 2^{-(2i+1)n^*}$. Picking $\rho$ uniformly from all restrictions with $|\rho^{-1}(*)| = v - 4\sqrt{n}$, is equivalent to picking randomly $T = \rho^{-1}(1) \cup \rho^{-1}(0)$ among all $4\sqrt{n}$-subsets of $V$, and then picking the restriction of $\rho$ to $T$ randomly among all vectors $x$ in $\{0, 1\}^{4\sqrt{n}}$. Let $k = \sqrt{n}/4$. Say $T$ is bad if

$$\mu \left( \left\{ x : \mu(Ext_{Q,T}(x)) \geq \frac{\alpha}{2k} \right\} \right) < \left( \frac{\alpha}{2} \right)^{2/k}$$

$T$ is good otherwise. We have

$$\Pr \left( \mu(Q_\rho) < \frac{\alpha}{2k} \right) \leq \Pr(T \text{ is bad})$$

$$+ \Pr \left( \mu(Q_\rho) < \frac{\alpha}{2k} \mid T \text{ is good} \right)$$

Note that $Q_\rho = Ext_{Q,T}(x)$. By the definition of goodness, and the choice of $k$, the second term is bounded by $1 - (\alpha/2)^{2/k} = o(1)$. Also note that $\alpha/2k \geq 2^{-(2i+1)n^*}$. It remains to bound the first term: We pick a random $T$ by first picking a random partition of $V$ into $4\sqrt{n}$-subsets and then picking a random subset from the partition. We must show that for
any partition, \( \Pr(T \text{ is bad}) < 1/2 \) for a random \( T \) in the partition. But this is precisely the content of corollary 4.4. 

Now we take care of G2:

Let \( A^* = \{ p \in A : \rho \leq \rho^{-1}(*) \} \). We say that \( \rho \) kills a path \( p_L \in A \) if there is no \( p_R \in Ext(p_L) \) with \( p_R \leq \rho^{-1}(1) \). We will show that for every choice of \( \rho \), \( \mu(A^*) \) is large and hence \( \Pr(\neg G2) \leq \Pr(\exists p_L \in A \text{ killed by } \rho) \).

**Claim 4.3** For every \( \rho \), \( \mu(A^*) \geq n^{-\varepsilon} \).

**Proof:** Recall that \( |\rho^{-1}(*)| = v - 4\sqrt{v} \) so that \( |V \setminus \rho^{-1}(*)| = 4\sqrt{v} \). It is easy to see that at most a fraction \( n^{-2/5} \) of the \( l' \)-subsets of \( V \) intersect \( V \setminus \rho^{-1}(*) \) so that at least a fraction \( 2n^{-\varepsilon} - n^{-2/5} \geq n^{-\varepsilon} \) of them are in \( A^* \).

We now use some combinatorics to bound the probability that there exists a path in \( A \) killed by \( \rho \):

**Claim 4.4** \( \Pr(\exists p_L \in A \text{ killed by } \rho) = o(1) \).

**Proof:** We have

\[
\Pr(\exists p_L \in A \text{ killed by } \rho) \\
\leq |A| \cdot \max_{p_L \in A} \{ \Pr(p_L \text{ is killed by } \rho) \}
\]

so let us look at the worst possible \( p_L \in A \). Note that \( supp(Ext(p_L)) \) contain at least a fraction \( n^{-2\varepsilon} \) of the \( l' \)-subsets of \( V \). This is because the worst situation is when all possible orderings of a given subset are contained in the collection.

Let \( F \equiv (\exists p_R \in supp(Ext(p_L)) \text{ with } p_R \leq \rho^{-1}(1)) \). We will bound now \( \Pr(F) \). We will break \( \Pr(F) \) into two cases according to how small \( |\rho^{-1}(1)| \) is. We overestimate \( \Pr(F) \) as follows:

\[
\Pr(F) \leq \Pr(|\rho^{-1}(1)| < \sqrt{v}) \\
+ \Pr(F \mid |\rho^{-1}(1)| \geq \sqrt{v})
\]

The first term can be bounded by \( \exp(\sqrt{v}/2) \) using the Chernoff bound. For the second term, for every random \( \rho \), we have \( |\rho^{-1}(1)| \geq \sqrt{v} \). By choosing a random \( \sqrt{v} \)-subset \( T \) of \( \rho^{-1}(1) \) we induce a marginal uniform distribution for \( T \) over all the \( \sqrt{v} \)-subsets of \( V \). We can now apply corollary 4.4 to get

\[
Pr(F \mid |\rho^{-1}(1)| \geq \sqrt{v}) \leq (1 - n^{-3\varepsilon})^{v^{1/2} - 1/10}
\]

so that \( \Pr(F) \leq \exp(-n^{1/4}) \). Recalling that \( |A| \) is less than \( n^{1/10} \), we conclude easily our calculations and get

\[
\Pr(\exists p \in A \text{ killed by } \rho) \leq \frac{n^{1/10} \cdot e^{-n^{1/4}}}{o(1)} = o(1)
\]

We have \( \Pr(\neg G1) + \Pr(\neg G2) \leq 1/2 + o(1) \) implying the existence of a good restriction. Take any consistent extension of each \( p_L \in A^* \) not killed by \( \rho \) to form \( \tilde{P} \). We have \( \mu(\tilde{P}) \geq n^{-\varepsilon} \) and lemma 4.4 is proved.

**ACKNOWLEDGMENTS**

We are very grateful to Miklos Ajtai for a conversation which led to the proof of theorem 4.1. We are indebted to Johan Hastad for allowing us to present his improvements of our early results.

**REFERENCES**