Probabilistic Communication Complexity of Boolean Relations

Ran Raz
Hebrew University

Avi Wigderson
Hebrew University

Abstract

In [KW] it was prooved that for every boolean function f there exist a communication complexity game R_f such that the minimal circuit-depth of f exactly equals to the communication complexity of R_f. If f is monotone then there also exists a game R_f^m with communication complexity exactly equals to the monotone depth of f. It was also prooved in [KW] that the communication complexity of $R^m_{st-connectivity}$ is $\Omega(\log^2 n)$, or eqivalently that the monotone depth of the s-t connectivity function is $\Omega(\log^2 n)$.

In this paper we consider the games R_f and R_f^m in a probabilistic model of communication complexity, and prove that the communication complexity of $R^m_{st-connectivity}$ is $\Omega(\log^2 n)$ even in the probabilistic case. We also prove that in every $NC1$ circuit for s-t connectivity at least a constant fraction of all input variables must be negated.

1 Introduction

In standard communication complexity, two players are trying to compute a function of their respective inputs, i.e. have to solve a decision problem. In [KW] (see also [K]), this model is extended to computing relations, i.e. search problems. The main motivation for the new model is that for any boolean function f, the deterministic communication complexity of a related relation (called here boolean relation) is exactly the circuit depth complexity of f.

[0]This research was partially supported by the American-Israeli Binational Science Foundation grant number 87-00082

Once a deterministic model exists, it is by now a reflex to compare it with various probabilistic (and non-deterministic) analogues. in this paper we indeed follow that reflex and discover some interesting results.

Every boolean relation has probabilistic communication complexity $O(\log n)$. This may seem discouraging, as most functions have depth $\Omega(n)$, and so the probabilistic model has nothing to do with circuit complexity ?!

In fact the real motivation is different. The new communication model seemed like a very promising tool for many reasons:

- The lack of computation.
- The ability to understand circuits top-down, and see wires as carriers of pairs of inputs, not of boolean values.
- The ability to define distributions on inputs that affect the communication complexity measure (but clearly not a rigid measure like circuit depth), and reveal the information theoretic reasons for lower bounds.
- Karchmer's thesis [K] gives many examples where both upper and lower bounds become clearer and more elegant when viewed through the eyes of the new model.
- It enables us to use powerful results in classical communication complexity. The best example is the recent lower bound for the monotone depth of the matching function [RW].
- It provided the intuition that enabled the proof of the $\Omega((\log n)^2)$ monotone depth bound for st-connectivity.

This paper tries to supply a deeper understanding of the st-connectivity monotone depth lower bound.The main result of this paper is an analogous $\Omega((\log n)^2)$ lower bound on the the probabilistic communication complexity of the relation associated with monotone st-connectivity.

A second result is a strange lower bound on the depth of non-monotone circuits for st-connectivity. In particular it will be proved that in every NC^1 circuit for st-connectivity at least a constant fraction of all input variables

must be negated. In general we will prove a trade-of between the depth and the number of negated input variables.

The paper is organized as follow:

The first section, which is an introduction, review the definitions of deterministic and probabilistic communication complexity, defines the probabilistic communication complexity of boolean relation, survey some interesting results about this subject and gives an introduction and notations for the proofs in the other sections.

The second section proves some usefull lemmas on which the other sections are based.

The third section gives the proof of the main result of the work. In order to make the proof clearer we left the proofs of some claims to the last section.

Section four gives the result about a non monotone circuit computing the st-connectivity function.

1.1 Deterministic Communication Complexity

Let X, Y, Z be finite sets, and $R \subset X \times Y \times Z$ be a relation.
For $x \in X, y \in Y$, let $Z(x,y) = \{z | (x,y,z) \in R\}$. The pairs (x,y) for which $Z(x,y) \neq \phi$ will be called the **edges** of the relation R, and denoted $E(R)$.

The communication problem for R is as follows. Two players, I and II get inputs $x \in X$ and $y \in Y$ respectively, with $(x,y) \in E(R)$. Their task is to **agree** on $z \in Z(x,y)$. The **communication complexity** of R, denoted $C(R)$, is the minimum number of bits, exchanged by the best protocol on a worst case input.

If $g : X \times Y \to Z$ is a function, then defining $R \subset X \times Y \times Z$ by $Z(x,y) = \{g(x,y)\}$ gives us the standard, well-studied communication problem.

The study of relations (search problems) was initiated by [KW], who showed that the circuit depth of a boolean function is exactly captured by the communication complexity of a related relation.

Specifically, let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function, and define the

two relations
$R_f, R_f^m \subset f^{-1}(1) \times f^{-1}(0) \times [n]$ by
$(x,y,i) \in R_f$ iff $x_i \neq y_i$, $(x,y,i) \in R_f^m$ iff $x_i = 1, y_i = 0$. In words, I gets an input $x \in \{0,1\}^n$ with $f(x) = 1$, II gets $y \in \{0,1\}^n$ with $f(y) = 0$, and their task in R_f (respectively R_f^m) is to find a coordinate where their inputs differ (respectively, a coordinate where x has 1 and y has 0).

Consider boolean circuits over the basis $\{\wedge, \vee, \neg\}$, where $\{\wedge, \vee\}$-gates have fan-in 2 and $\neg$-gates are applied to inputs only. A circuit is monotone if it contains no negations. Let $d(f)$, (respectively, $d^m(f)$) denote the minimum depth of a circuit (respectivly monotone circuit) computing f. Then,

Theorem 1.1 (KW) *:*

1. *For every boolean function f, $d(f) = C(R_f)$).*
2. *For every monotone boolean function f, $d^m(f) = C(R_f^m)$.*

The main result of [KW] was to use this equivalence to prove a lower bound on the monotone depth of the function st-connectivity. This function, $stc : \{0,1\}^n \to \{0,1\}$ interprets the input as the edge set of an undirected graph G on vertex set V with $s, t \in V$, ($n = \begin{pmatrix} |V| \\ 2 \end{pmatrix}$), and tests if s and t are connected in G.

Theorem 1.2 (KW) $d_{stc}^m = C(R_{stc}^m) = \Omega(\log^2 n)$.

We will remark that if the function f is monotone the comunnication problem R_f^m is equivalent to a problem in which player I has a minterm x of f, player II has a maxterm y of f and $(x,y,i) \in R_f^m$ iff $x_i = 1, y_i = 0$. For example in the st-connectivity problem R_f^m can be presented as a problem in which player I has a st-path , player II has a st-coloring, and their goal is to find a bi-colored edge in the path.

1.2 Probabilistic Communication Complexity

Let $R \subset X \times Y \times Z$. A **deterministic protocol** P for R is a pair of algorithms (one for each player), according to which each of them sends messages (depending on its input and previous messages), whose final output is an element $P(x,y) \in Z$. A protocol is correct if
$\forall (x,y) \in E(R), P(x,y) \in Z(x,y)$. Let $\mathcal{P}(R)$ denote the set of correct deterministic protocols for R.

A **probabilistic protocol** P allows the algorithms of the players to be probabilistic, i.e. depend also on the tosses of random coins. The complexity of a protocol P on input $(x,y) \in E(R)$ is the expected number (over the coin flips) of bits sent, denoted $c_P(x,y)$. The complexity of P, denoted $c(P)$ is $\max_{(x,y)\in R} c_P(x,y)$.

We distinguish between two probabilistic models, one called COMMON, in which both players share a **common** random bit string, and one called PRIVATE, in which each player has a **private** random string. The PRIVATE model is clearly weaker and more realistic than COMMON.

Note that $P(x,y)$ becomes a random variable, and we are interested in the event $P(x,y) \in Z(x,y)$.

For $0 \leq \epsilon < 1/2$ define,

$\mathcal{P}_\epsilon(R) = \{P \in \text{COMMON} | \forall (x,y) \in E(R), Pr[P(x,y) \notin Z(x,y)] \leq \epsilon\}$,

$\tilde{\mathcal{P}}_\epsilon(R) = \{P \in \text{PRIVATE} | \forall (x,y) \in E(R), Pr[P(x,y) \notin Z(x,y)] \leq \epsilon\}$,

and the complexity of private and common protocols by

$C_\epsilon(R) = \min_{P \in \mathcal{P}_\epsilon(R)} c(P)$,

$\tilde{\mathcal{C}}_\epsilon(R) = \min_{P \in \tilde{\mathcal{P}}_\epsilon(R)} c(P)$.

Also note that $C(R) = \min_{P \in \mathcal{P}(R)} c(P)$.

Some obvious facts about the above measures are that for every relation R:

1. $\tilde{\mathcal{P}}_\epsilon(R) \subset \mathcal{P}_\epsilon(R)$, and so $C_\epsilon(R) \leq \tilde{\mathcal{C}}_\epsilon(R)$.

2. For $\delta < \epsilon < 1/2$, $C_\epsilon(R) \leq C_\delta(R)$, $\tilde{\mathcal{C}}_\epsilon(R) \leq \tilde{\mathcal{C}}_\delta(R)$.

3. Protocols in $\mathcal{P}_0(R)$ always answer correctly (Las Vegas protocols), and hence $\mathcal{P}_0(R)$ consists precisely of convex combinations of protocols in $\mathcal{P}(R)$, and in particular $C_0(R) \leq C(R)$.

A simple but fundemental observation of Newman allows us to deal with the simple COMMON model only. It saies that COMMON is stronger than PRIVATE by at most $O(\log n)$ bits.

Theorem 1.3 (N) *For every $\delta > \epsilon > 0$ or $\delta = \epsilon = 0$ $\tilde{C}_\delta(R) \leq C_\epsilon(R) + O(\log n)$ where $n = \log(|X||Y|)$*

Proof: (sketch) A COMMON protocol for R is actually a set G of protocols. The two players use their random string to choose a protocol from G and apply this protocol on their input pair. For each input pair the probability in G for correct answer is at least $1 - \epsilon$. By the Chernoff bound, there exists a subset $\tilde{G} \subset G$ of size $O(\log(|X||Y|))$ such that for each input pair the probability in $\tilde{G}$ for correct answer is at least $1 - \delta$. We can now define a new PRIVATE protocol, in which player I chooses a protocol in $\tilde{G}$ and tells his choice to the second player ($\log|\tilde{G}|$ bits), the two players apply this protocol on their input pair, to get an answer which must be correct by probability of at least $1 - \delta$. We can choose $\tilde{G}$ in such a way that the complexity will not increase by much.

1.3 Probabilistic Communication Complexity of Boolean Relations

We now return to relations of the form R_f of boolean functions $f : \{0,1\}^n \to \{0,1\}$. We also consider two "universal" relations $U, U^m \subset \{0,1\}^n \times \{0,1\}^n \times [n]$, defined by:

$(x, y, i) \in U$ iff $x_i \neq y_i$

$(x, y, i) \in U^m$ iff $x_i = 1$ and $y_i = 0$.

The universal relation U (respectively U^m) captures the simultaneous computation of all (respectively all monotone) n-bit functions. In fact, a

deterministic protocol for U (U^m) yields a universal formula (e.g. see [W]) for all n-bit functions (monotone functions).

We call the relations R_f, R_f^m, U, U^m **boolean relations**. For boolean relations, bounded error protocols are not much stronger than Las Vegas protocols.

Lemma 1.1 *Let R be a boolean relation. Then*

$C_0(R) \leq (C_\epsilon(R)+2)(\log_{1/\epsilon} n + 1)$

$\tilde{\mathcal{C}}_0(R) \leq (\tilde{\mathcal{C}}_\epsilon(R)+2)(\log_{1/\epsilon} n + 1)$

Proof: (sketch) The correctness of the answer a protocol gives is easily verified by each player, at the cost of one additional bit. The players repeat a bounded error protocol $\log_{1/\epsilon} n$ times. If all failed (an event of probability $\leq 1/n$), the players exchange their inputs.

The following theorem, du to Hastad, shows that in both the COMMON and PRIVATE models the universal relation U (and hence every relation R_f) has small probabilistic complexity.

Theorem 1.4 (H) $\tilde{\mathcal{C}}_0(U) = C_0(U) = O(\log n)$.

By theorem 1.3 it is enough to give a proof for the COMMON model. We will describe here a proof of Karchmer.

Proof (K1) (sketch) Notice that $(x,y) \in E(U)$ simply means $x \neq y$. Think of the common random string as $\log n$ boolean n-bit vectors $\{a^i\}$. Player I sends the sequence of bits $\{b^i =< a^i, x >\}$ [1] to player II. With probability $\geq 1 - 1/n$, $b^i \neq < a^i, y >$ for some i. Let S be the set of positions where a^i is 1. The players construct $\tilde{x}, \tilde{y} \in \{0,1\}^S$ by dropping the other coordinates from their vectors. As $\tilde{x}$,$\tilde{y}$ have diffrent parities, an $O(\log n)$ protocol for parity finds a position where $\tilde{x}$ differs from $\tilde{y}$. If $b^i =< a^i, y >$ for every i , the players exchange their inputs.

This theorem provides an exponential gap between probabilistic and deterministic communication complexity. It therefore shows that probabilistic communication complexity is not directly related to non-monotone depth

[1] $< \alpha, \beta >$ denotes the inner product mod 2

complexity of functions, as most of them have depth $\Omega(n)$. The next theorem provides an exponential gap between the probabilistic monotone and non-monotone games.

Theorem 1.5 $C_\epsilon(U^m) = \Omega(n)$.

Proof Note that if player *II* complements its input vector y, and both players regard their inputs as subsets of $[n]$ indexed by 1's, then $U^m \subset 2^{[n]} \times 2^{[n]} \times [n]$, with $(S,T,z) \in U^m$ iff $z \in S \cap T$. Moreover, $(S,T) \in E(U^m)$ iff $S \cap T \neq \phi$. Recall that the disjointness function, $D \subset 2^{[n]} \times 2^{[n]} \times \{0,1\}$ is defined by $(S,T,1) \in D$ iff $S \cap T \neq \phi$. But since the answer of a protocol for U^m can be easily verified, a trivial reduction from D to U^m gives $C_\epsilon(D) \leq O(C_\epsilon(U^m))$, and the remarkable lower bound of [KS], $C_\epsilon(D) = \Omega(n)$, gives the result.

We conclude this sub-section remarking that [RW], using a more subtle reduction to disjointness, were recently able to prove an exponential gap between the probabilistic monotone and non-monotone games, for the specific relation associated with the perfect matching function.

1.4 Probabilistic Protocols for $s-t$-connectivity

As in [KW], we view the inputs of the players in R^m_{stc} as an $s-t$ path for player *I*, and $s-t$ cut for player *II*. Let $R^m_{stc(l)}$ be the restriction of R^m_{stc} when the $s-t$ path of player *I* has length l. In [KW] it was proved:

Theorem 1.6 (KW) *For every* $l \leq n^{1/10}$, $C(R^m_{stc(l)}) = \Omega(\log n \cdot \log l)$.

The main theorem of this work is an analogous result for probabilistic computation.

Theorem 1.7 $\forall \epsilon < 1/2, l \leq n^{1/1000}, C_\epsilon(R^m_{stc(l)}) = \Omega(\log n \cdot \log l)$.

Corollary 1.1 $C_\epsilon(R^m_{stc}) = \Omega((\log n)^2)$.

The full proof of theorem 1.7 appears in third section. We just say a few words about it here. We consider pairs *(path,cut)* of inputs to the players for which the relation is a function, i.e. when the path crosses the cut exactly once. Intuitively, theses are the difficult inputs, and we prove that no "shallow" protocol can give **the** correct answer on most of them (this implies the probabilistic lower bound).

At a very high level, we follow the lines of the inductive proof of [KW]. The main difficulties arise for two reasons. First, on top of the obvious parameters, we have to bound the quality of the algorithm, i.e. the fraction of errors. Second, the set of inputs we consider is **not** a cartesian product of inputs to each player.

As the proof of [KW] is our starting point, familiarity with it will ease understanding of this proof.

1.5 Notations

Let N be a set of vertices, $|N| = n$. Let s, t be special vertices in N.
Define:
A_N = the set of all paths of length l in N starting at s, and ending at t.
B_N = the set of all bi-partitions of N, for which s and t are in different parts.
We will think of B_N's elements also as colorings.
For a subset $U \subset N$ we will use the notations:
A_U = the set of all paths of length l in $U \bigcup \{s, t\}$ starting at s, and ending at t.
B_U = the set of all two-colorings of U by the colors: s and t.
Define:
$O^i_N = \{(a, b) | a \in A_N, b \in B_N$. The i'th edge in the path a is the only edge 2-colorable according to the coloring $b\}$
$O_N = \bigcup_i O^i_N$.
For $A \subset A_N, B \subset B_N$ define:
$O^i_{AB} = \{(a, b) | a \in A, b \in B, (a, b) \in O^i_N\}$
$O_{AB} = \bigcup_i O^i_{AB} = \{(a, b) | a \in A, b \in B, (a, b) \in O_N\}$
For $a \in A_N, B \subset B_N$ define:
$O^i_{aB} = \{(a, b) | b \in B, (a, b) \in O^i_N\} = O^i_{\{a\}B}$

$O_{aB} = \bigcup_i O^i_{aB} = \{(a,b) | b \in B, (a,b) \in O_N\} = O_{\{a\}B}$

For $A \subset A_N, b \in B_N$ define:

$O^i_{Ab} = \{(a,b) | a \in A, (a,b) \in O^i_N\} = O^i_{A\{b\}}$

$O_{Ab} = \bigcup_i O^i_{Ab} = \{(a,b) | a \in A, (a,b) \in O_N\} = O_{A\{b\}}$.

Let $P : S \to R$ be a probability distribution on a finite set S. Let $H(P)$ be the entropy of P, and $I(P) = \log_2 |S| - H(P)$ be the information of P.

In all following we will use asymptotic notation, and the claims will hold for n sufficiently large.

2 Useful Lemmas

The first lemma shows that when the information is small, one gets good bounds on the probability of events with known support.

Lemma 2.1 *Let S be a finite set, $|S| = s$. Let $P : S \to R$ be a distribution on S, where the information is known to be $I(P) \leq \theta$. Let $w < 1/100$ and assume $\sqrt{\frac{4\theta}{w}} < 1/10$. Define*

$\bar{a} = \sum P(x)$, where the sum is taken over the sw largest elements, and

$\underline{a} = \sum P(x)$, where the sum is taken over the sw smallest elements. Then

$$\bar{a} \leq w\left(1 + \sqrt{\frac{4\theta}{w}}\right), \underline{a} \geq w\left(1 - \sqrt{\frac{4\theta}{w}}\right).$$

Proof: First, let us prove the claim for $\bar{a}$. Let us redistribute $P(x)$'s mass so that it will be equal on the sw largest elements, and also equal on all the others. This cannot change the value of $\bar{a}$, but decreases $I(P)$. Assume therefore that

$$P(x) = \begin{cases} \frac{1}{s}(1+\Delta) & \text{for the } ws \text{ largest elements} \\ \frac{1}{s}\left(1 - \frac{\Delta w}{1-w}\right) & \text{otherwise} \end{cases}$$

and then:

$$\theta \geq I(P) = -w\log(1+\Delta) - (1-w)\log\left(1 - \frac{\Delta w}{1-w}\right)$$

Expanding $\log(1+\Delta)$ to its Taylor series, we get for small (c_1, c_2):

$\theta \geq -w\Delta + \frac{w\Delta^2}{2} + \frac{c_1}{6}w\Delta^3 + w\Delta + \frac{\Delta^2 w^2}{2(1-w)} + \frac{c_2}{6}w^3\Delta^3$. If we assume $\Delta < 1/10$, then $|c_1|, |c_2| < 2$ and we get $\theta \geq I(P) \geq w\Delta^2/4$, which implies $\Delta \leq \sqrt{\frac{4\theta}{w}}$, i.e. a much stronger upper bound on Δ follows. This fact and the fact that $I(p)$ is monotone in Δ evident that the assumption $\Delta < 1/10$ was correct.

The bound on $\underline{a}$ is proved similarly.
End of proof of lemma 2.1.

Lemma 2.2 states that the bipartite graph O^i_N is very regular in the sense of Szemeredi [S], i.e. every subgraph that is not too small has, upto a very small error, the expected number of edges.

Lemma 2.2 *Let $l \leq n^{1/100}$. Let*
$A \subset A_N$ be a set for which $|A| = \alpha|A_N| \geq n^{-1/100}|A_N|$, and let
$B \subset B_N$ be a set for which $|B| = \beta|B_N| \geq 2^{-n^{1/100}}|B_N|$. Then for all i,
$\alpha\beta|O^i_N|(1 - O(n^{-1/10})) \leq |O^i_{AB}| \leq \alpha\beta|O^i_N|(1 + O(n^{-1/10}))$.

Proof: The idea of the proof is to project O^i_{AB} on all small subsets of N. On "most" of them the projection is "well behaved", and we can use lemma 2.1 to bound $|O^i_{AB}|$ by averaging over these projections.

For a set $Q \subset B_N$ define P_Q to be a probability measure on B_N which is uniformly distributed on Q :
$P_Q(b) = \begin{cases} 1/|Q| & b \in Q \\ 0 & \text{otherwise} \end{cases}$
For $s, t \in U \subset N$, define $P_{Q/U}$ to be the projection of P_Q on B_U.
For $s, t \in U \subset N$ and for a subset $L \subset A_N$, define $L_{/U} = L \bigcap A_U$, which is the set of all paths in L that are contained in U.
Let us consider all the subsets $U \subset N$ for which $s, t \in U$,
and $|U| = n^{1/2-1/100}$.
For every i define:
$\gamma^i_U = \frac{|O^i_{A_{/U},B}|}{|O^i_{A_U,B_N}|}$, $\gamma^i = \frac{|O^i_{AB}|}{|O^i_N|}$.
Intuitively, γ^i = the fraction of pairs in O^i_N which exist in our system, while γ^i_U = the fraction of pairs in the system for which the path is contained in U.

In this notation, lemma 2.2 states:
$\alpha\beta(1-O(n^{-1/10})) \leq \gamma^i \leq \alpha\beta(1+O(n^{-1/10}))$
Note that $|O^i_{A_U,B_N}|$ is fixed for all the U's as $|U|$ is fixed.
Therefore $\gamma^i = \overline{\gamma^i_U}$, where the bar denotes averaging over all possible U's.
Let us bound this average.
$$|O^i_{A_{/U},B}| = \sum_{b\in B}|O^i_{A_{/U},b}| = \sum_{b'\in B_U}|B|P_{B/U}(b')|O^i_{A_{/U},b'}| = |B|\sum_{b'\in B_U}P_{B/U}(b')|O^i_{A_{/U},b'}|,$$
$$|O^i_{A_U,B_N}| = \sum_{b\in B_N}|O^i_{A_U,b}| = \sum_{b'\in B_U}|B_N|P_{B_N/U}(b')|O^i_{A_U,b'}| = |B_N|\sum_{b'\in B_U}P_{B_N/U}(b')|O^i_{A_U,b'}|.$$
Both of the last sums are subdivided into two sums according to the hamming weight of the partition b', which we will denote by $h(b')$. We will say that b' is "usual" if
$|U|/2 - |U|^{3/4} \leq h(b') \leq |U|/2 + |U|^{3/4}$, and otherwise we will call b' "unusual". We get:
$$|O^i_{A_{/U},B}| = |B|\sum_{b'usual}P_{B/U}(b')|O^i_{A_{/U},b'}| + |B|\sum_{b'unusual}P_{B/U}(b')|O^i_{A_{/U},b'}|$$
$$|O^i_{A_U,B_N}| = |B_N|\sum_{b'usual}P_{B_N/U}(b')|O^i_{A_U,b'}| + |B_N|\sum_{b'unusual}P_{B_N/U}(b')|O^i_{A_U,b'}|.$$

Claim 2.1 *The "unusual" part in both equations is $o(n^{-1/10})$ of the "usual" part in both formulas, and hence can be omitted.*

The next claim states that for "usual" b', $|O^i_{A_U,b'}|$ is almost independent of b'.

Claim 2.2 *For some K, depending only on n, l, i, and for every "usual" b', $|O^i_{A_U,b'}| = K(1+o(n^{-1/9}))$.*

Combine claims 2.1 and 2.2 to estimate γ^i_U (upto $(1+O(n^{-1/10}))$). $\gamma^i_U = \dfrac{|O^i_{A_{/U},B}|}{|O^i_{A_U,B_N}|} \approx \dfrac{|B|}{|B_N|}\dfrac{\sum_{b'usual}P_{B/U}(b')|O^i_{A_{/U},b'}|}{\sum_{b'usual}P_{B_N/U}(b')|O^i_{A_U,b'}|} \approx \beta\left(\sum_{b'usual}P_{B/U}(b')\dfrac{|O^i_{A_{/U},b'}|}{K}\right)$.

Let $\alpha_U = |A_{/U}|/|A_U|$, the density of paths in U. By summing over all possible b''s, we get: $\dfrac{\sum_{b'}|O^i_{A_{/U},b'}|}{K} \approx |B_U|\alpha_U$.
(= up to a factor of $1+O(n^{-\frac{1}{9}})$)

Looking at the expression for γ_U^i we see that all the $P_{B/U}(b')$ are weighted in the sum, and the sum of the weights $\approx |B_U|\alpha_U$. As all the weights are in $[0,1]$, we notice:
$\beta\sum_{Y_1} P_{B/U}(b') \leq \gamma_U^i \leq \beta\sum_{Y_2} P_{B/U}(b')$,
(up to a factor of $1+O(n^{-\frac{1}{9}})$)
where Y_1 are the $\alpha_U|B_U|$ smallest elements of $P_{B/U}(b')$, and Y_2 are the $\alpha_U|B_U|$ largest elements of $P_{B/U}(b')$.

The lemma's claim is about γ^i, where we know that $\gamma^i = \overline{\gamma_U^i}$. Define a new probability function: $p(b',U) = P_{B/U}(b')/N_U$, defined for pairs $(U, b' \in B_U)$, where N_U denotes the number of possible sets U. Then:

$\overline{\gamma_U^i} \leq \overline{\beta \sum P_{B/U}(b')} \leq \beta \sum p(b',U)$,
where the first sum is taken over the $\alpha_U|B_U|$ largest elements for each U, and the second sum is taken over the $\alpha|B_U|N_U$ largest elements $p(b',U)$. This bound on γ_U^i is correct as $\overline{\alpha_U} = \alpha$ and $p(b',U)$ contains all the distributions $P_{B/U}$, and we allow taking more of the largest elements of one of the distributions.

Similarly, we can prove a lower bound, and combining the two bounds we have:
$\beta \sum p(b',U) \leq \overline{\gamma_U^i} \leq \beta \sum p(b',U)$, where the first sum is over the $\alpha|B_U|N_U$ smallest elements and the second over the $\alpha|B_U|N_U$ largest elements.

Claim 2.3 *The information* $\overline{I(P_{B/U})} \leq n^{-1/2}$.

As $I(p(b',U)) = \overline{I(P_{B/U})} \leq n^{-1/2}$, and using lemma 2.1 we have:
$$\alpha\beta(1-O(n^{-1/10})) \leq \alpha\beta\left(1-\sqrt{\frac{4n^{-1/2}}{\alpha}}\right) \leq \overline{\gamma_U^i} \leq \alpha\beta\left(1+\sqrt{\frac{4n^{-1/2}}{\alpha}}\right)$$
$$\leq \alpha\beta(1+O(n^{-1/10}))$$
End of proof of lemma 2.2.

The next corollary follows immediately after observing that $|O_N^i|$ is independent of i.

Corollary 2.1 $|O_{AB}^i|$ *is independent of* i *up to a factor of* $1+O(n^{-1/10})$.

Lemma 2.3 states that if we start with a family of colorings B with high enough density in B_N , and consider the subset of B that agrees with a random coloring I of a small random set U , the density of this subset in B_{N-U} is likely to be close to the one of B in B_N. A similar lemma was used in [KW].

Lemma 2.3 *Let $B \subset B_N$, where $|B| = \beta|B_N| \geq 2^{-n^{1/100}}|B_N|$.*

Choose $U \subset N - \{s\} - \{t\}$, where $|U| = n^\gamma$, randomly with uniform probability on all such sets. Assume $\gamma = 1/2$. Color U by a random coloring I, and denote by $B_{/I}$ the set of colorings in B which agree with I on U. Then with probability $\geq 1 - O(n^{-1/10})$, we have
$\beta|B_{N-U}|(1 - n^{-1/10}) \leq |B_{/I}| \leq \beta|B_{N-U}|(1 + n^{-1/10})$

Proof: Let N_U be the number of sets U possible. On the probability space $\{(b,U)|b \in B_U\}$, define a probability function $p(b,U) = \frac{P_{B/U}(b)}{N_U}$ as before. Then we get
$\overline{I(P_{B/U})} \leq n^{\gamma - 99/100} = n^{-49/100}$. Therefore
$I(p(b,U)) = \overline{I(P_{B/U})} \leq n^{\gamma-99/100} = n^{-49/100} \stackrel{\text{def}}{=} \theta$.

A coloring I will not be within the lemma's bounds if:

1. $p(I,U) \geq \frac{1}{N_U|B_U|}(1 + n^{-1/10})$, or
2. $p(I,U) \leq \frac{1}{N_U|B_U|}(1 - n^{-1/10})$.

Let $\Delta = n^{-1/10}$, w_1 be the probability that the first condition holds for (I,U), and w_2 the probability that the second condition holds for (I,U). Then, using the inequalities derived in lemma 2.1, we get
$\Delta \leq \sqrt{\frac{4\theta}{w_1}}$, $\Delta \leq \sqrt{\frac{4\theta}{w_2}}$, and therefore
$w_1 + w_2 \leq \frac{8\theta}{\Delta^2} \approx n^{-29/100} \leq n^{-1/10}$.
End of proof of lemma 2.3.

Lemma 2.4 generalizes lemma 2.3, and points to a key distinction between this proof and the one of [KW]. Here lemma 2.3 does not suffice, as we shall have to consider a **different** family of colorings per path.

Lemma 2.4 *Let M be a set of vertices for which $|M| = m$. Let $q < m^{1/100}$, and E_M be the set of all subsets of M with cardinality q. Let $E \subset E_M$ be a set with $|E| = \delta|E_M|$, for some $\delta > m^{-1/100}$. For every $e \in E$, let C_e be a set of two-colorings on $M - e$ (so $C_e \subset B_{M-e}$). Assume that for every e $\beta_e \stackrel{\text{def}}{=} |C_e|/|B_{M-e}| > \beta > 2^{-m^{1/100}}$.*

Choose randomly and uniformly a subset $U \subset M$ with cardinality $|U| = m^{1/2}$, and color it randomly and uniformly with a coloring I. Let U_s, U_t be the parts of U colored as s, t respectively. Then with probability $\geq 1/4$, there exists $e \in E$ satisfying:

1. $e \subset U_t$
2. $\beta_e|B_{M-U}|(1 - m^{-1/10}) \leq |C_{e/I}| \leq \beta_e|B_{M-U}|(1 + m^{-1/10})$.

($C_{e/I}$ is the set of colorings in C_e which agree with I on U.)

Proof of lemma 2.4: The lemma seems very easy because with very high probability there is e for which 1 hold, and for each e 2 hold with very high probability (lemma 2.3). Unfortunately this is not enough (as some examples show) the problem is the dependence between the choice of e and the event 2. We force independence by refining the probability space.

First choose U_t as an **ordered** set, and then choose U_s as an **un-ordered** set. Let f be the order on U_t. With exponentially high probability, $|U_t| > 1/4\sqrt{m}$. Divide the first $\sqrt{m}/4$ (by f) points of U_t into $k = \sqrt{m}/4q$ sets of q elements each, which will be named $V_1, V_2, \ldots, V_k$. With exponentially high probability there exists an i for which $V_i \in E$, so we will assume such an i exists, and let i_0 be the smallest such i. Let $e = V_{i_0}$, and $p_i = Pr(i = i_0)$.

Let i' be the smallest i for which
$\sum_{i=1}^{i'} p_i = Pr(i_0 \leq i') > 1/2$.

To each triplet (U_s, U_t, f) we match the pair (e, i_0), and for each (e, i_0) we will define $T(e, i_0)$ to be the inverse image of (e, i_0) (e.g. all the triplets

corresponding to (e, i_0)). Clearly, if (e, i) was chosen, by definition $e \in E$, and condition (1) is satisfied.

Claim 2.4 *If a pair (e, i) was obtained as above with $i \leq i'$, then e satisfies condition (2) with probability $> 1/2$.*

Using this claim and the fact that the pair (e, i_0) satisfies $i_0 \leq i'$ with probability $> 1/2$ we see that the e we defined satisfies lemma 2.4 with probability $> 1/4$.
End of proof of lemma 2.4.

3 Main Theorems and Corollaries

As in [KW], we gauge the parameters: N (number of vertices), l (path length), α (density of "existing" paths A), and β (densitity of existing colorings B). However, here the algorithm makes mistakes, and we keep another parameter γ,which is the density of correct answers for pairs in O_{AB}.

As in [KW], we show how to significantly increase the density α without affecting β by much, after assigning random colors to a small subset of vertices, and decreasing l. However, to harness the decrease in γ and l, we have to maintain bounds on $\gamma^{100}l$ and on $\alpha\gamma$.

In the proof of the theorem we are going to have two cases:
In the first case we will increase α to $O(\alpha^{0.5})$ without affecting β and $\gamma^{100}l$ by mutch. In the second case we will increase $\gamma^{100}l$ without affecting α and β by mutch.

If case 1 hold we got an increase in α. Otherwise case 2 hold and we increase $\gamma^{100}l$ again and again antil γ is very big. When γ is big enough case 1 must hold.

Theorem 3.1 *For $l < n^{1/100}$, let $A \subset A_N$, $|A| = \alpha|A_N|$ for some $\alpha > n^{-1/1000}$, and $B \subset B_N$, $|B| = \beta|B_N|$ for some $\beta > 2^{-n^{1/1000}}$. Assume*

there exists a communication protocol which when given $a \in A, b \in B$, produces after τ communication bits some edge in the graph which is a candidate to be an edge in the path a which is 2-colored according to b. Define:
$T = \{(a,b) | (a,b) \in O_{AB}$, *and the algorithm is correct on* $(a,b)\}$.
Let $\gamma = |T|/|O_{AB}|$, *and assume that for some fixed small* ϵ, $\gamma^{100} > n^{-\epsilon}$. *Then on a set* M *with cardinality* $m = n - O(\sqrt{n}\log(n))$ *and vertices* $s, t \in M$, *for some* $\tilde{l}$, *we have* $\tilde{A} \subset A_M$, $\tilde{B} \subset B_M$ *such that:*
$|\tilde{A}| = \tilde{\alpha}|A_M|$, $|\tilde{B}| = \tilde{\beta}|B_M|$. *And there exists a communication protocol which when given* $\tilde{a} \in \tilde{A}, \tilde{b} \in \tilde{B}$, *produces after* τ *communication bits an edge which is a candidate for being an edge in* $\tilde{a}$ *that is bi-colored according to* $\tilde{b}$. *And if*
$\tilde{T} = \{(\tilde{a},\tilde{b}) | (\tilde{a},\tilde{b}) \in O_{\tilde{A},\tilde{B}}$, *and the algorithm is correct on* $(\tilde{a},\tilde{b})\}$
and $\tilde{\gamma} = |\tilde{T}|/|O_{\tilde{A},\tilde{B}}|$, *then*

1. $\tilde{\alpha} > n^{-\delta(\epsilon)} K_\alpha \sqrt{\alpha}$.
2. $\tilde{\beta} > K_\beta \beta$.
3. $\tilde{l}\tilde{\gamma}^{100} > K_\gamma l \gamma^{100}$.

$K_\alpha, K_\beta, K_\gamma$ *are some fixed constants and* $\delta(\epsilon)$ *satisfies* $\lim_{\epsilon \to 0} \delta(\epsilon) = 0$.

We will use the following notation: $T^i = T \bigcap O^i_{AB}, \gamma^i = |T^i|/|O^i_{AB}|$.
For $C \subset A, D \subset B$, define:
$T^i_{CD} = T^i \bigcap O^i_{CD}$,
$\gamma^i_{CD} = |T^i_{CD}|/|O^i_{CD}|$,
$T_{CD} = T \bigcap O_{CD}$,
$\gamma_{CD} = |T_{CD}|/|O_{CD}|$,
and for $D \subset B, a \in A$ define
$T^i_{aD} = T^i \bigcap O^i_{aD}$,
$\gamma^i_{aD} = |T^i_{aD}|/|O^i_{aD}|$,
$T_{aD} = T \bigcap O_{aD}$,
$\gamma_{aD} = |T_{aD}|/|O_{aD}|$

Proof of the theorem : First we split the path so as to roughly balance the product $\gamma^{100} l$ on each part. Look at the partitions of the path of length l into a left part of length l_1, and a right part of length l_2 ($l = l_1 + l_2$).

Define: $\gamma_1 = |\bigcup_{i=1}^{l_1} T^i|/|\bigcup_{i=1}^{l_1} O^i_{AB}|$, $\gamma_2 = |\bigcup_{i=l_1+1}^{l} T^i|/|\bigcup_{i=l_1+1}^{l} O^i_{AB}|$
(γ_1, γ_2 are the weighted averages of γ^i on both parts of the path respectively).

Claim 3.1 *There exists a choice for l_1, l_2 satisfying*

1. $\gamma_1, \gamma_2 \geq (1/2 - o(1))\gamma$
2. $\gamma_1^{100} l_1, \gamma_2^{100} l_2 \geq (1/2 - o(1))\gamma^{100} l$

From this point on l_1, l_2 are fixed by this choice.

Next we observe the quality of the algorithm on colorings that correspond to fixed paths on each part. Let
$A_1 = \{a \in A : |\bigcup_{i=1}^{l_1} T^i_{aB}|/|\bigcup_{i=1}^{l_1} O^i_{aB_N}| > \beta\gamma_1/10\}$
$A_2 = \{a \in A : |\bigcup_{i=l_1+1}^{l} T^i_{aB}|/|\bigcup_{i=l_1+1}^{l} O^i_{aB_N}| > \beta\gamma_2/10\}$
Intuitivly A_1,A_2 are the sets of paths in A in which there is high quality in the left side or in the right side (respectively).
We will divide the proof into two parts:

CASE 1: $|A_1| > 3|A|/4$, $|A_2| > 3|A|/4$.

This is the simpler case. Here, the algorithm is of high quality, in both sides, on many paths. As in [KW], paths in one of the parts are much denser, and an application of lemma 2.4 performs the reduction.

Define $A' = A_1 \bigcap A_2$. Then $|A'| > |A|/2$. Let $\alpha' = |A'|/|A_N|$ (the density of $|A'|$). Let C be the set of all paths on N of length l_1 starting at the special vertex s. Let D be the set of all paths of length l_2 ending at t. Then every path $a \in A_N$ is combined of a path $c \in C$ and a path $d \in D$. Let us use the notations $a = c \times d$, $A_N = C \times D$. Define
$A_C = \{c \in C : |\{d \in D : c \times d \in A'\}| > \alpha'/4\}$,
$A_D = \{d \in D : |\{c \in C : c \times d \in A'\}| > \alpha'/4\}$,
$\alpha_C = |A_C|/|C|, \alpha_D = |A_D|/|D|$.
Then clearly either $\alpha_C > \sqrt{\alpha'/2}$ or $\alpha_D > \sqrt{\alpha'/2}$. Assume w.l.o.g that

$\alpha_C > \sqrt{\alpha'/2}$. Consider subsets $U \subset N$ for which $|U| = \sqrt{n}$, and colorings I of U.

Definition:

$$A_{c,U,I} = \left\{ c \in A_C \;\middle|\; \begin{array}{ll} \text{i)} & c \bigcap U = \phi \\ \text{ii)} & \exists d \in D \text{ s.t. } a = c \times d \in A', d \subset U_t, \\ & |\bigcup_{i=1}^{l_1} T^i_{aB/I}| / |\bigcup_{i=1}^{l_1} O_{aB_{N-U}}| > \beta\gamma_1/20 \\ \text{iii)} & 2\beta|B_{N-U}| > |B_{/I}| > \beta|B_{N-U}|/2 \end{array} \right.$$

Intuitively the set $A_{c,U,I}$ is the set of all the left paths c for which $c \bigcap U = \phi$ and the following restriction on U, I will be sucsesfull, that mean : there will be a right path $d \subset U_t$ for which $c \times d$ is a high quality path in A', even when we consider only partitions which are consistent with I on U.

Claim 3.2 *there is U, I for which $|A_{c,U,I}|/|C| > \alpha_C/5$*

Let U, I be such that the conditions hold for them, Define: $M = N - U, m = n - n^{1/2}, \tilde{l} = l_1, \tilde{A} = A_{c,U,I}$ (after adding t to the end of each path in $A_{c,U,I}$), $\tilde{B} = B_{/I}$ (when $B_{/I}$ are considered as colorings of M). Define a communication protocol which uses τ communication bits, and for which the theorem holds, as follows:
For all $c \in \tilde{A}$ there exists a d_c which establishes condition ii) in the definition. Given $\tilde{a} \in \tilde{A}, \tilde{b} \in \tilde{B}$, the player holding the path will extend $\tilde{a}$ to $c \times d_c$ on the set $N = M \bigcup U$, where c is $\tilde{a}$ after removing t from its end. The player holding $\tilde{b}$ will extend it to b by adding the coloring I on the set U. Using the algorithm given in the theorem, they will get an edge in a which is a candidate for being bi-colored in b.

By claim 3.2 $\tilde{\alpha} > \alpha_C/5 > \sqrt{\alpha'}/10 > \sqrt{\alpha}/20$.

By condition iii) $\tilde{\beta} > \beta/2$. By condition ii) $\tilde{\gamma} > \gamma_1/20$.

Using claim 3.1,
$\gamma_1^{100}\tilde{l} = \gamma_1^{100} l_1 \geq \gamma^{100} l/2 \Rightarrow \tilde{\gamma}^{100}\tilde{l} > K_\gamma \gamma^{100} l$.
End of CASE 1.

CASE 2: Either $|A_1| < 3|A|/4$ or $|A_2| < 3|A|/4$.

Assume w.l.o.g that $|A_1| < 3|A|/4$, and let $\alpha_1 = |A_1|/|A_N|$.

In this case the algorithm's quality (γ_1) is high on relatively few paths (α_1) in the left part. However, there is a tradeoff between α_1 and γ_1. We will show that after a random restriction their product, and hence α, will not decrease too much. On the other hand, γ, and hence $\gamma^{100}l$ will increase by a constant factor. Repeating it enough times will force the conditions of Case 1 to hold, without losing too much in path density.

Let $f(a) \stackrel{\text{def}}{=} |\bigcup_{i=1}^{l_1} T^i_{aB}|/|\bigcup_{i=1}^{l_1} O^i_{aB_N}|$.

Intuitively $f(a)$ is the quality of the path a on the left side.

We will do a very strange thing and ommit all the paths a with a very high quality. This is done only in order to simplifie the computations. Define:

$A'_1 = \{a \in A_1 | f(a) < 2\beta\}$

We will show now that by restricting to A'_1 we did not loste a lot of "quality volume".

Claim 3.3 $\sum_{a \in A'_1} f(a) > \frac{9}{10}(1 - o(1))\beta\gamma_1|A|$

As $|A'_1| < 3|A|/4$, the average of $f(a)$ on A'_1 is greater than $1.2(1 - o(1))\beta\gamma_1$. Let $\alpha'_1 = |A'_1|/|A_N|$, then $\alpha'_1\overline{f(a)} > \frac{9}{10}(1-o(1))\beta\gamma_1\alpha$, where $\overline{f(a)}$ denotes the average over A'_1.

The next step is to pass to a subset A' of A'_1 in which every left path of one path is a left path of a lot of paths. We will show that by passing to A' we do not loste a lot of "quality volume":

Define the sets C, D as in the Case 1 (C is the set of all paths on N of length l_1 starting at the special vertex s. D is the set of all paths of length l_2 ending at t. Then every path $a \in A_N$ is combined of a path $c \in C$ and a path $d \in D$. we are using the notation $a = c \times d$)

Define:

$A_C = \{c \in C : |\{d \in D : c \times d \in A'_1\}| > |D|\gamma_1\alpha'_1/1000\}$,

$A' = \{(c \times d) \in A'_1 | c \in A_C\}$.

Then $\alpha' = |A'|/|A_N| > (1 - \gamma_1/1000)\alpha'_1$.

As for all $a \in A'_1$, $f(a) < 2\beta$, we also have $\sum_{a \in A'_1 - A'} f(a) < \frac{2}{1000}\beta\gamma_1\alpha'_1|A_N|$.

From this and claim 3.3 follows:

Claim 3.4 *In A' we have:*

1. $\alpha'\overline{f(a)} > \frac{9}{10}(1-o(1))\beta\gamma_1\alpha$
2. $\overline{f(a)} > 1.2(1-o(1))\beta\gamma_1$ *(2 is following from 1 and the size of A')*

These two bounds (1 and 2) we are going to carry with us to the end of the proof. We will use bound 2 which intuitively saies that the quality is bigger than what we could expect (by a factor of at least 1.2).
For $c \in A_C$, define $\mu(c) = |\{d \in D : c \times d \in A'\}|/|D|$,
$f(c) = \sum_{c\times d \in A'} \mu(c) f(c \times d)/|D|$
Intuitively $\mu(c)$ is the measure of right-paths d that can be matched to the left-path c. $f(c)$ is the quality of the left-path c.
We want now to show that the two bounds 1 and 2 hold also for a subset $\tilde{A}_C \subset A_C$

Claim 3.5 *there exists $\tilde{A}_C \subset A_C, |\tilde{A}_C|/|C| = \alpha'$, in which the following holds for the average $\overline{f(c)}$:*

1. $\alpha'\overline{f(c)} > \frac{9}{10}(1-o(1))\beta\gamma_1\alpha$
2. $\overline{f(c)} > 1.2(1-o(1))\beta\gamma_1$

For $c \in \tilde{A}_C$, let $D_c = \{d \in D | c \times d \in A', f(c \times d) \geq \frac{24}{25} f(c)\}$.
To facilitate the use of lemma 2.4, we prove

Claim 3.6 *We can assume that for every $c \in \tilde{A}_C, |D_c| \geq n^{-1/100}|D|$*

Just as in Case 1, we get that for a random choice of U, I, for every $c \in \tilde{A}_C$, there is a probability of at least $1/5$ that:

1. $c \bigcap U = \phi$
2. $\exists d \in D_c, d \subset U_t, |\bigcup_{i=1}^{l_1} T^i_{c\times d, B/I}|/|\bigcup_{i=1}^{l_1} O^i_{aB_{N-U}}| > \frac{11.5}{12} f(c)$
3. $\beta|B_{N-U}|(1-O(n^{-1/10})) < |B_{/I}| < \beta|B_{N-U}|(1+O(n^{-1/10}))$

For all U, I, let $A_{(U,I)}$ be the set of all $c \in \tilde{A}_C$ for which the three conditions hold for (U, I). The set $A_{(U,I)}$ is the set of all the good left-paths according to reduction by (U, I). Every c appears in at least a $1/5$ of the $A_{(U,I)}$, and we will arbitrary remove it from some of the sets so it appears in exactly $1/5$ of them. Let $f_{(U,I)} = \sum_{c \in A_{(U,I)}} f(c)$. Intuitively $f_{(U,I)}$ is the quality of the reduction by (U, I).

Claim 3.7 *There exists a (U, I) for which ...*

(a). $f_{(U,I)} > \frac{1}{500}\alpha' \overline{f(c)}|C|$

(b). $f_{(U,I)}/|A_{(U,I)}| > \frac{11}{11.5}\overline{f(c)}$, *(where $\overline{f(c)}$ is averaged over $\tilde{A}_C$).*

Take (U, I) for which the claim holds. For each $c \in A_{(U,I)}$, let d_c be the d for which condition (2) holds (the condition after claim 3.6). Define a new algorithm exactly as in Case (1), where:
$M = N - U, m = n - n^{1/2}, \tilde{l} = l_1, \tilde{A} = A_{(U,I)}$ (after adding t to each path), and $\tilde{B} = B_{/I}$.

Combining second conditions of claims 3.7 and 3.5 and that $\tilde{\gamma} > \frac{11.5}{12} f(U, I)/\beta|A_{(U,I)}|$ (by condition 2) , we get $\tilde{\gamma} > \frac{11}{10}\gamma_1$, and therefore:

1. $\tilde{\gamma}^{100}\tilde{l} > 10\gamma_1^{100} l_1 > 4\gamma^{100} l$
   Combining first conditions of claims 3.7 and 3.5 and that $\tilde{\alpha} = |A_{(U,I)}|/|C|$, $\tilde{\gamma} > \frac{11.5}{12} f(U, I)/\beta|A_{(U,I)}|$, and that $\gamma_1 > \frac{1}{2}\gamma$ we get:
2. $\tilde{\alpha}\tilde{\gamma} > \frac{1}{2000}\alpha\gamma$
   As nothing changes β by much we also get:
3. $\tilde{\beta} > (1 - O(n^{-1/10}))\beta$.

End of CASE 2.

To finish the proof, notice that we can iterate the procedure described in the second case, until the first case's conditions hold. We will iterate at most $\epsilon \log n$ as $\gamma^{100} l$ increases at each iteration by 4 and cannot be larger than l (because $\gamma < 1$),and because at the beginning $\gamma^{100} > n^{-\epsilon}$. When γ is big enough Case 1 must hold. During the iterations β decreases by $(1 - o(1))$,

$\gamma^{100} l$ increases, and $\alpha\gamma$ decreases by $(\frac{1}{2000})^{\epsilon \log n} = n^{-\delta(\epsilon)}$ at most. So for the final α, which we denote by $\tilde{\alpha}$,, we get: $\tilde{\alpha} > \alpha n^{-\delta(\epsilon)} \frac{\gamma}{\tilde{\gamma}} > \alpha n^{-\delta(\epsilon)-\epsilon} = \alpha n^{-\delta'(\epsilon)}$. Now we apply the Case 1, and finally
$\tilde{\beta} > K_\beta \beta$
$\tilde{\alpha} > K_\alpha n^{-\delta'(\epsilon)} \sqrt{\alpha}$
$\tilde{l}\tilde{\gamma}^{100} > K_\gamma l \gamma^{100}$
End of proof of theorem .

As in [KW], it is useful (and proves a stronger statement) to consider structured protocols.

Definition 3.1 *A structured protocol works in stages. In each stage the player holding the path sends $c_1 \log n$ bits, and the player holding the coloring answers with n^{c_1} bits. Here we assume $c_1 < 1/10000$.*

Theorem 3.2 *For $l \leq n^{1/100}$, assume there exists a structured communication protocol which when given $a \in A, b \in B$, produces an edge in the path a which is candidate for being 2-colorable according to b. Define as before:*
$T = \{(a,b) | (a,b) \in O_{AB},$ *the algorithm is correct on* $(a,b)\}$,
$\gamma = |T|/|O_{AB}|, \alpha = |A|/|A_N|, \beta = |B|/|B_N|$, *and assume* $\gamma^{100} \geq n^{-\epsilon}$ *(ϵ from theorem 3.1),* $\alpha \geq n^{-1/2000}, \beta \geq 2^{-n^{1/2000}}$. *Assume the algorithm has τ steps. Then there exists a set M of cardinality $m = n - O(\sqrt{n}\log(n))$, and $\tilde{l}$, so that*
$\tilde{A} \subset A_M, |\tilde{A}| = \tilde{\alpha}|A_M|$
$\tilde{B} \subset B_M, |\tilde{B}| = \tilde{\beta}|B_M|$,
and there exists a structured communication algorithm of $\tau - 1$ stages. Such that if
$\tilde{T} = \{(\tilde{a},\tilde{b}) | (\tilde{a},\tilde{b}) \in O_{\tilde{A},\tilde{B}},$ *the algorithm is correct on*$(\tilde{a},\tilde{b})\}$, *and*
$\tilde{\gamma} = |\tilde{T}|/|O_{\tilde{A}\tilde{B}}|$, *then the following holds:*
$\tilde{\alpha} \geq n^{-1/2000}, \tilde{\beta} \geq K_\beta \beta 2^{-2n^{c_1}}$ *(the K_β from theorem 3.1),*
and $\tilde{l}\tilde{\gamma}^{100} \geq K'_\gamma l \gamma^{100}$.

Proof of theorem 3.2: Observe the first step of the algorithm. The $c_1 \log n$ bits sent by the first player induce a partition of A into $k = n^{c_1}$ parts $A_1, \ldots, A_k$, and giving information about which part a belongs to. For each $j \leq k$, let:

$\alpha_j = |A_j|/|A_N|$, and $\gamma_j = |T_{A_j,B}|/|O_{A,B}|$, where $T_{A_j,B}$ is defined as for theorem 3.1 .

Claim 3.8 *$\exists j$ for which*

1. $\alpha_j \geq n^{-1/1500}$
2. $\gamma_j^{100} \geq \gamma^{100}/2$

Let us now look at A_{j_0} for which the claim holds, and look at the branch of the algorithm from A_{j_0} (i.e, assume $a \in A_{j_0}$).

Observe the partition of B into $k = 2^{n^{c_1}}$ parts $B_1, \ldots, B_k$. Let $\beta_j = |B_j|/|B_N|, \gamma_j = |T_{A_{j_0}B_j}|/|O_{A_{j_0}B_j}|$. As before, we get j_1 for which $\beta_{j_1} \geq \beta 2^{-2n^{c_1}}$, $\gamma_{j_1}^{100} \geq \gamma_{j_0}^{100}/2 \geq \gamma^{100}/4$. We got sets A_{j_0}, B_{j_0} for which $\alpha_{j_0} \geq n^{-1/1500}$, $\beta_{j_0} \geq \beta 2^{-2n^{c_1}}$. Continuing the algorithm for $\tau - 1$ more steps produces an algorithm for which $l\gamma_{j_1}^{100}$ is smaller than the original $l\gamma^{100}$ by a factor of 4 at most. Applying theorem 3.1 on this algorithm concludes the proof of this theorem.

Corollary 3.1 *For $l \leq n^{1/100}$, there is no structured communication algorithm of $\tau = c_2 \log l$ stages (for some constant c_2), which when given $a \in A_N, b \in B_N$ produces an edge in a which is a candidate for being 2-colored according to b, and is right on half of the pairs of $O_{A_N B_N}$.*

Proof: Assume such an algorithm exists. By using theorem 3.2 repeatedly, we get a sequence of algorithms. For the last of them the following holds:
$\tilde{\alpha} \geq n^{-1/2000}$
$\tilde{\beta} \geq (2^{-2n_1^{c_1}} K_\beta)^\tau = 2^{-2n_1^{c_1} c_2 \log l} K_\beta^\tau \geq 2^{-n_1^{2c_1}} \geq 2^{-n_1^{2/10000}}$
$\tilde{l}\tilde{\gamma}^{100} \geq l(\frac{1}{2})^{100}(\frac{1}{k_\gamma})^{c_2 \log l} \geq \sqrt{l}$.
(The condition about γ^{100} ($\gamma^{100} > n^{-\epsilon}$) holds in each iteration because in every iteration γ^{100} decreases by a constant (at most), and because c_1 is arbitrary small.)
This algorithm is supposed to work in 0 steps, i.e. without passing any information, which is impossible as for the remaining $\tilde{A}, \tilde{B}$ the conditions of lemma 2.2 still hold, and $\tilde{\gamma} > n^{-\epsilon}$ for a very small ϵ.

Corollary 3.2 *For $l \leq n^{1/100}$ there is no structured probabilistic communication algorithm of $\tau = c_1 \log l$ which gives for every pair (a,b) acorrect answer with probability 1/2.*

A simple argument [KW] shows that each stage of a structured protocol can simulate $\Omega(\log n)$ bits of any protocol. hence we obtain our main result, theorem 1.7 in:

Corollary 3.3 *$\forall \epsilon < 1/2, l \leq n^{1/100}$ we have $C_\epsilon(R^m_{stc(l)}) = \Omega(\log n \log l)$. In particular, $C_\epsilon(R^m_{stc}) = \Omega((\log n)^2)$.*

4 The Number Of Negated Variables In S-T Connectivity

This section is devoted to a trade-of between the depth and the number of negated variables in a circuit computing the *st* -connectivity problem. The proof we are going to give is based on the lower bound which was proved in section 3. The result does not necessarily depend on this lower bound, however the intuition came from it.

To state this lower bound, recall that we deal with circuits in which negations are applied only to input variables. Since *s*—*t* connectivity is monotone, every input variable must occur positively, and if none occur negated, the depth is $\Omega((\log n)^2)$. Suppose there is a shallower non-monotone circuit. How many inputs should be negated?

Theorem 4.1 *$\exists \alpha > 0$, so that $\forall l$, every circuit for s-t connectivity with depth $\leq \alpha \log n \log l$ has $\Omega((n/l)^2)$ negated input variables. (Here n is the number of vertices, so there are $\binom{n}{2}$ input variables.)*

Corollary 4.1 *Every NC^1 circuit for stc has $\Omega(n^2)$ negated inputs.*

In order to proove the theorem we will define a new problem:
The ST-connectivity problem, will be the one in which $S, T \subset N$ with $|S| =$

$|T| = n/3$ and the problem is to say wheter there is a path between S and T.

Notice that a minterm in the new problem is a path between S and T which use only vertices in $N - (S \bigcup T)$. A maxterm is a st-coloring of $N - (S \bigcup T)$. The problems of st-connectivity and ST-connectivity are related by two simple reductions: The first reduction from st-connectivity to ST-connectivity shows that the same lower bounds which we have for the depth of monotone st-connectivity and for the probabilistic communication complexity of the related relation, are also good for the ST-connectivity problem. The reduction is simply to create S and T from n new vertices each, and to connect s to S and t to T.
The second reduction from ST-connectivity to st-connectivity shows that it is enough to proove the promised trade-off for the problem of ST-connectivity. The reduction is simply to add two new vertices : s and t ,and to connect s to all the vertices in S and t to all the vertices in T.

Proof of theorem 4.1:

Assume C is a shallow circuit (of depth $\leq \alpha \log n \log l$) for ST-connectivity with $O((n/l)^2)$ negated variables. Let this set of variables be $E \subset \begin{pmatrix} [n] \\ 2 \end{pmatrix}$. Now C can be thought of as a protocol, such that given a graph G_1, containing an ST-path to player I, and G_2 with no ST-path to player II, returns either an edge in $G_1 - G_2$, or an edge in $(G_2 - G_1) \bigcap E$.

We probabilistically reduce the monotone relation R^m_{ST} (the relation related to monotone ST-connectivity) to C. Player I gets an ST-path P, and player II a cut Q. They interpret the common random string as a permutation $\pi : N \to N$ in which $\pi(T) = T$ and $\pi(S) = S$ and respectively construct $G_1 =$ edges in $\pi(P)$, $G_2 =$ (edges in $\pi(Q)) - E$. Now they apply C, and get a response in $G_1 - G_2$ (they do not get a response in $(G_2 - G_1) \bigcap E$ because $G_2 \bigcap E$ is empty). Inspection shows that the response must be a double-colored edge if $\pi(P) \bigcap E = \phi$.

A straightforward calculation shows that for small enough α the probability for $\pi(P) \bigcap E \neq \phi$ is half at most. Therefore, in such a probability we got a double-colored edge ,but this contradicts theorem 1.7.

5 Proofs of Claims

Proof of claim 2.1:

First we will show that the "unusual" part in the second equation is negligible compared to the "usual" part:
$|B_N| \sum_{b' unusual} P_{B_N/U}(b')|O^i_{A_U,b'}| = \frac{|B_N|}{|B_U|} \sum_{b' unusual} |O^i_{A_U,b'}| < \frac{|B_N|}{|B_U|} \sum_{b' unusual} |U|^l =$
$\frac{|B_N||U|^l}{|B_U|} \sum_{b' unusual} 1 < \frac{|B_N||U|^{(l+1)}}{|B_U|} \left(\begin{array}{c} |U| \\ |U|/2 + |U|^{3/4} \end{array} \right)$.

Using elementary approximations, we get:
$\left(\begin{array}{c} s \\ s/2 + s^{3/4} \end{array} \right) \approx 2^s \left(\frac{1}{e}\right)^{2s^{1/2}}$. Hence the "unusual" summand in the previous sum is $< \frac{|B_N|}{|B_U|} 2^s 2^{-s^{1/2}}$, while the "usual" summand is $> \frac{|B_N|}{|B_U|} 2^s \cdot 1$, and therefore their ratio is $\geq 2^{-s^{1/2}} \approx 2^{-n^{1/4}}$.

As we have to prove that on average
$|O^i_{A_{/U},B}| = |O^i_{A_U,B_N}|\alpha\beta(1 + O(n^{-1/10}))$, and as $|O^i_{A_U,B_N}|$ is fixed for all U, we have that a term which is multiplied by a factor of $(1 + O(2^{-n^{1/4}}))$ is insignificant, and therefore we can ignore the right hand term in $|O^i_{A_U,B_N}|$.

Now we consider the equation on $|O^i_{A_{/U},B}|$. The "unusual" term in the sum $|O^i_{A_{/U},B}|$ is smaller than the one in $|O^i_{A_U,B_N}|$. Therefore as $\alpha\beta \geq n^{-1/100}2^{-n^{1/100}} >> 2^{-n^{1/4}}$, we have that the "unusual" term is negligible also in $|O^i_{A_{/U},B}|$.
End of proof of claim 2.1.

Proof of Claim 2.2:

$|O^i_{A_U,b'}| = \left(\begin{array}{c} h(b') \\ l-i \end{array} \right) \left(\begin{array}{c} |U| - h(b') \\ i \end{array} \right) (l-i)!i!$, and therefore
$(h(b') - (l-i))^{l-i}(|U| - h(b') - i)^i < |O^i_{A_U,b'}| < h(b')^{l-i}(|U| - h(b'))^i$.
We write $h(b') = |U|/2 + c|U|^{3/4}$, for some $-1 \leq c \leq 1$, and then both the upper and lower bound are equal to:
$(|U|/2 + c'|U|^{3/4})^{l-i}(|U|/2 - c''|U|^{3/4})^i$,
where $-2 < c', c'' < 2$, and the terms $i, l-i$, were included in c', c''. Therefore

the bounds are equal to:
$(|U|/2)^l \left(1+\frac{c'}{|U|^{1/4}}\right)^{l-i}\left(1-\frac{c''}{|U|^{1/4}}\right)^{i}$. Put $K=(|U|/2|)^l$. Then the dependency on b' is only in the term
$I=\left(1+\frac{c'}{|U|^{1/4}}\right)^{l-i}\left(1-\frac{c''}{|U|^{1/4}}\right)^{i}$, but as $|U|^{1/4}\approx n^{1/8}$, and as $l\leq n^{1/100}$, we have $I=1+o(n^{-1/9})$.

End of proof of claim 2.2.

Proof of claim 2.3:

The entropy $H(P_B)=\log_2|B|\geq|B_N|-n^{-1/100}=n-n^{-1/100}$.
It is known from information theory that
$\overline{H(P_{B/U})}\geq\frac{H(P_B)|U|}{n}$ and therefore
$\overline{H(P_{B/U})}\geq|U|-\frac{|U|}{n^{99/100}}$, or
$\overline{I(P_{B/U})}\leq|U|n^{-99/100}=n^{-1/2}$.

End of proof of claim 2.3.

Proof of claim 2.4: $T(e,i)$ contains all triplets (U_s,U_t,f) for which

1. $e=V_i$
2. There is no $j<i$ for which $V_j\in E$.

$P(\exists j<i, V_j\in E)=p(i_0<i)\leq p(i_0<i')\leq 1/2$, and therefore 2 occurs with probability $\geq 1/2$, and $T(e,i)$ contains at least half of the triplets for which $e=V_i$ (because 1,2 are independed events). Corresponding to each triplet is also the projection pair (U_s,U_t) which is a different representation of the pair (U,I). Consider the pairs (U,I) for which $e\subset U_t$. Let such a pair be called a "good" pair if
$\beta_e|B_{M-U}|(1-m^{-1/10})\leq|C_{e/I}|\leq\beta_e|B_{M-U}|(1+m^{-1/10})$, and otherwise call

the pair "bad". By lemma 2.3, the fraction of pairs which are "bad" is $< O(m^{-1/10})$. Returning to the corresponding triplets, we get that the fraction of "bad" triplets (triplets whose projection is "bad") is $< O(m^{-1/10})$. As $T(e,i)$ contains more than half the triplets for which $e = V_i$, the probability of a triplet in $T(e,i)$ to be "bad" is $< 2O(m^{-1/10}) < 1/2$. Therefore the probability that e satisfies the lemma is $> 1/2$.

End of proof of claim 2.4.

Proof of claim 3.1

Part 1 of the claim follows easily from part 2, which we prove below.

From corollary 2.1, we have:

$$\gamma_1 = \left(\sum_{i=1}^{l_1}\gamma^i/l_1\right)(1+O(n^{-1/10})), \gamma_2 = \left(\sum_{i=l_1+1}^{l}\gamma^i/l_2\right)(1+O(n^{-1/10})),$$

$$\gamma = \left(\sum_{i=1}^{l}\gamma^i/l\right)(1+O(n^{-1/10})), \text{ and therefore}$$

$$(1+O(n^{-1/10}))\gamma l = \gamma_1 l_1 + \gamma_2 l_2.$$

By convexity of x^{100}, and because $l = l_1 + l_2$ we get:

$(1+O(n^{-\frac{1}{10}})\gamma^{100} l \leq \gamma_1^{100} l_1 + \gamma_2^{100} l_2$

But for every i $\gamma^i \leq 1$ and therefore by the bound we have on $\gamma^{100} l$ there exists a choice for l_1,l_2 s.t. $\gamma_1^{100} l_1, \gamma_2^{100} l_2 \geq (1/2 - o(1))\gamma^{100} l$.

End of proof of claim 3.1.

Proof of claim 3.2:

For $c \in A_C$ choose randomly and uniformly U, I. By lemma 2.4, with probability $> 1/4$ there exists $d \in D$ for which condition ii) holds. This is by using in lemma 2.4 $M = N - l_1$, $q = l_2 < m^{1/100}$, and the set

$E = \{d \in D | c \times d \in A'\}$, where a path corresponds to a subset containing exactly the path's vertices. Then by definition of A_C, $|E|/|E_N| > \alpha'/4$. For each $e \in E$ the lemma's set C_e will be the set of colorings in B for which e

is colored with color t, which does not agree with c on a single edge, and the algorithm is correct on the pair $((c \times e), coloring)$ (i.e. C_e is the set of right components in $\bigcup_{i=1}^{l_1} T^i_{c\times e,B}$). Since $c \times e \in A' \subset A_1$, and by A_1's definition, we get in lemma 2.4:

$|B_{M-e}|\beta_e > \frac{1}{10}\gamma_1\beta|\bigcup_{i=1}^{l_1} O^i_{aB_N}| = \frac{1}{10}\gamma_1\beta|B_{N-l}| \geq \frac{1}{10}\gamma_1\beta|B_{M-e}|$

So by the bounds we have on β ,γ we know that β_e is big enough.

This means that for all $c \in A_C$, and for random choice of U, I, the second condition in the definition holds with probability $> 1/4$. Condition i) holds with probability $1 - o(1)$ as $|U||c| < n^{3/4}$. Condition iii) holds with probability of at least $1 - O(n^{-1/10})$ by lemma 2.3. Therefore with probability of at least $1/5$, all 3 conditions hold, and there exist U, I for which $|A_{c,U,I}|/|A_C| > 1/5$, or $|A_{c,U,I,}|/|C| > \alpha_C/5$.

End of proof of claim 3.2.

Proof of claim 3.3:

In the set $A - A_1$, by definition of A_1 $f(a) < \beta\gamma_1/10$. $|A - A_1| < |A|$, and therefore:

1. $\sum_{a\in A-A_1} f(a) < \beta\gamma_1|A|/10$.
   The average of $f(a)$ on the set A is $\beta\gamma_1$ up to a factor of $(1+O(n^{-1/10}))$
   (This is true by definition of γ_1, by the fact that $|\bigcup_{i=1}^{l_1} O^i_{aB_N}|$ is constant with a and because $|O^i_{AB}|/|O^i_{AB_N}| = \beta(1 + O(n^{-\frac{1}{10}}))$) therefore:
2. $\sum_{a\in A} f(a) \approx \beta\gamma_1|A|$
   Let $Z = \{a \in A_1 | f(a) > 2\beta\}$. For all $a \in Z$
   $2\beta < f(a) < |\bigcup_{i=1}^{l_1} O^i_{aB}|/|\bigcup_{i=1}^{l_1} O^i_{aB_N}|$. Summing on the elements of Z:

$|\bigcup_{i=1}^{l_1} O^i_{ZB}| > 2\beta|\bigcup_{i=1}^{l_1} O^i_{A_N B_N}|\frac{|Z|}{|A_N|}$, and therefore Z is too small for satisfying the conditions in lemma 2.2. So $|Z| < n^{-1/100}|A_N|$, and also

3. $\sum_{a\in Z} f(a) < 2\beta n^{-1/100}|A_N|$ (otherwise we will increase Z to a set of cardinality $n^{-1/100}|A_N|$, which would contradict lemma 2.2).

We have $A'_1 = A_1 - Z$. $\sum_{a\in A} f(a) = \sum_{a\in A-A_1} f(a) + \sum_{a\in A'_1} f(a) + \sum_{a\in Z} f(a)$, and by (1),(2),(3):

$$\sum_{a\in A'_1} f(a) > (1 - O(n^{-1/10}))\beta\gamma_1|A| - \beta\gamma_1|A|/10 - 2\beta n^{-1/100}|A_N|.$$

As $\gamma_1|A| >> n^{-1/100}|A_N|$, we have

$$\sum_{a\in A'_1} f(a) > \frac{9}{10}(1 - o(1))\beta\gamma_1|A|.$$

End of proof of claim 3.3.

Proof of claim 3.5: By bound 1 of 3.4 $\alpha' \sum_{a\in A'} f(a) > \frac{9}{10}(1-o(1))\beta\gamma_1\alpha|A'|$.

But $\sum_{a\in A'} f(a) = \sum_{c\in C}\mu(c)f(c)|D|$ (when $c \notin A_C$, $\mu(c) = f(c) = 0$).

By definitions of α' and $\mu(c)$ we know that the sum of weights $\sum_{c\in C}\mu(c) = \alpha'|C|$. So taking $\tilde{A}_C$ to be the $\alpha'|C|$ largest elements:

$|D|\sum_{\tilde{A}_C} f(c) \geq |D|\sum_{c\in C}\mu(c)f(c) = \sum_{a\in A'} f(a)$,

If $\overline{f(c)}$ denotes the average over $\tilde{A}_C$, we have:

$\overline{f(c)} = \sum_{\tilde{A}_C} f(c)/|\tilde{A}_C| = \sum_{\tilde{A}_C} f(c)/|C|\alpha' \geq \sum_{A'} f(a)/|D||C|\alpha'$, and using claim 3.4(1), $\alpha'\overline{f(c)} > \frac{9}{10}(1 - o(1))\beta\gamma_1\alpha$.

(2) follows from (1) and the bound on α'/α.

End of proof of claim 3.5.

Proof of claim 3.6: By the fact that $f(c)$ is the average of $f(c \times d)$ on A' and by the fact that $f(c \times d) \leq 2\beta$ we get: $|D_c| \geq \frac{1}{25} f(c)\mu(c)|D|/2\beta$. From $\tilde{A}_C$, discard all the c's for which $f(c) < n^{-1/1000}\beta\gamma_1$, we will be left with a set $\tilde{A}'_C$ which satisfies the conditions of claim 3.5 up to small constants . Therefore, assume $\tilde{A}_C = \tilde{A}'_C$, and substituting in the previous inequality: $f(c) \geq n^{-\frac{1}{1000}}\beta\gamma_1$, $\mu(c) \geq \frac{1}{1000}\gamma_1\alpha'_1$ (definition of A_C), we get $|D_c| \geq Kn^{-1/1000}\gamma_1^2\alpha'_1|D|$. But, $\alpha'\overline{f(a)} > \frac{9}{10}(1+o(1))\beta\gamma_1\alpha$, $\overline{f(a)} < 2\beta$ and $\alpha'_1 > \alpha'$, (the averaging is over A') .Putting it all together, we get: $\alpha'_1 > K'\gamma_1\alpha$. So, $|D_c| \geq \tilde{K}n^{-1/1000}\gamma_1^3\alpha|D| \Rightarrow |D_c| > n^{-1/100}|D|$.
(α, γ_1 are big enough)

End of proof of claim 3.6.

Proof of claim 3.7:

Let χ be the number of possible (U, I). Define:
$W = \{(c, U, I) | c \in A_{(U,I)}\}$. Then
$\sum_W f(c) = \frac{1}{5}\alpha'\overline{f(c)}|C|\chi$. Define:
$W' = \{(c, U, I) \in W |$ (a) holds for$(U, I)\}$. Then
$\sum_{W-W'} f(c) < \frac{1}{500}\alpha'\overline{f(c)}|C|\chi$. Therefore on W',
$\sum_{W'} f(c) > (\frac{1}{5} - \frac{1}{500})\alpha'\overline{f(c)}|C|\chi$, and as
$|W'| < |W| = \frac{1}{5}\alpha'|C|\chi$, we have that the average of $f(c)$ on W' is at least $\overline{f(c)}(1 - \frac{1}{100})$. Therefore there is at least one (U, I) for which condition (a) holds and for which the average of $f(c)$ is at least $\overline{f(c)}(1 - \frac{1}{100})$, which means that condition (b) holds.

End of proof of claim 3.7.

Proof of claim 3.8: $\sum_{j=1}^{k}\alpha_j = \alpha$. As $T_{AB} = \bigcup T_{A_j,B}$, we have $\sum_{j=1}^{k}\gamma_j|O_{A_j,B}| = \gamma|O_{AB}|$. Partition the sum into two sums : over the sets A_j with cardinality $\alpha_j \geq n^{-1/1500}$, and over the sets with smaller cardinality. Now $\sum_{\alpha_j<n^{-1/1500}} \gamma_j|O_{A_jB}| \leq \sum_{\alpha_j<n^{-1/1500}} |O_{A_j,B}| = |O_{A'B}|$, where $A' = \bigcup_{\alpha_j<n^{-1/1500}} A_j$. $|O_{A'B}| \leq 2n^{-1/100000}\alpha\beta|O_{A_NB_N}|$, follows from lemma 2.2 because

$$|A'| < (n^{-1/1500}n^{1/10000}/\alpha)\alpha|A_N| < n^{-1/100000}\alpha|A_N|$$

, and if the inequality does not hold, we get a contradiction to lemma 2.2 by increasing A' until its cardinality is $n^{-1/100000}\alpha|A_N|$. Therefore:
$\sum_{\alpha_j<n^{-1/1500}} \gamma_j|O_{A_jB}| \leq 2n^{-1/100000}\alpha\beta|O_{A_NB_N}| \leq n^{-1/200000}\gamma|O_{AB}|$.
(This is true because by lemma 2.2 $\alpha\beta|O_{A_NB_N}| \approx |O_{AB}|$, and because $\gamma \geq n^{-\epsilon/100}$ for a very small ϵ)
By using lemma 2.2, if $\alpha_j \geq n^{-\frac{1}{1500}}$ then $\gamma_j|O_{A_jB}| = \gamma_j\alpha_j\beta|O_{A_NB_N}|(1 + O(n^{-1/10}))$, and therefore by substituting in $\sum_{j=1}^{k}\gamma_j|O_{A_j,B}| = \gamma|O_{AB}|$. we get $\sum_{\alpha_j\geq n^{-1/1500}} \alpha_j\gamma_j = \alpha\gamma(1 + O(n^{-1/200000}))$. By convexity of $y = x^{100}$, and because $\sum\alpha_j \approx \alpha$ (up to $O(n^{-1/200000})$, we have
$\sum_{\alpha_j\geq n^{-1/1500}} \alpha_j\gamma_j^{100} > \alpha\gamma^{100}(1 + O(n^{-1/200000})) > \alpha\gamma^{100}/2$. So, there exists a j for which $\alpha_j \geq n^{-1/1500}$, and $\gamma_j^{100} > \gamma^{100}/2$.

End of Proof of claim 3.8.

ACKNOWLEDGEMENTS

We are greatful to Paul Beame for a careful reading of this manuscript and discovering an error in the proof of claim 3.8.

BIBLIOGRAPHY

[**CK**] I. Csiszár and J. Körner "Information Theory - Coding Theorems for Discrete Memoryless Systems", *Academic Press*, NY 1982.

[**H**] J. Hastad, Private Communication.

[**K**] M. Karchmer "The Complexity of computation and restricted Machines" *Ph.D Thesis, The Hebrew University,* 1988.

[**K1**] M. Karchmer, Private Communication.

[**KS**] B. Kalyanasundaram and G. Snitger "The Probabilistic Communication Complexity of Set Intersection", *Proceedings Structure in Complexity Theorey* pp.41-49, 1987.

[**KW**] M. Karchmer and A. Wigderson "Monotone Circuits for Connectivity Require Super-logarithmic Depth" *Proceedings of the 20th STOC*, pp. 539-550, 1988.

[**N**] I. Newman, Private Communication.

[**RW**] R. Raz and A. Wigderson "Monotone Circuits for Matching Require Linear Depth" *Proceedings of the 22th STOC*, pp. 287-292, 1990.

[**S**] E. Szemeredi "Regular Partitions of Graphs", *Problems combinatories et theorie les graphs, Coll. CNRS* pp. 399-401, 1976.

[**W**] I. Wegner, *The Complexity of Boolean Functions* John Wiley, 1988.

[**Y**] A. C.-C. Yao, "Some Complexity Questions Related to Distributive Computing", *Proceedings of* 11^{th} *STOC*, pp. 209-213 (1979).