Electronic Colloquium on Computational Complexity, Report No. 96 (2025)

# Searching for Falsified Clause in Random $(\log n)$-CNFs is Hard for Randomized Communication

Artur Riazanov
*EPFL*
artur.riazanov@epfl.ch

Anastasia Sofronova
*EPFL*
anastasiia.sofronova@epfl.ch

Dmitry Sokolov
*EPFL*
sokolov.dmt@gmail.com

Weiqiang Yuan
*EPFL*
weiqiang.yuan@epfl.ch

July 16, 2025

**Abstract**

We show that for a randomly sampled unsatisfiable $O(\log n)$-CNF over $n$ variables the randomized two-party communication cost of finding a clause falsified by the given variable assignment is linear in $n$.

## 1 Introduction

This paper studies the communication complexity of *Falsified Clause Search Problem*.

**Definition 1 ([LNNW95]).** Let $X, Y$ be two disjoint sets of boolean variables and $\varphi$ be a CNF formula over the variables $X \sqcup Y$. We define *Falsified Clause Search Problem* or $\mathrm{Search}_\varphi$ associated with formula $\varphi$ in the following way:

*input:* a pair $(x, y) \in \{0,1\}^X \times \{0,1\}^Y$;

*output:* a clause $C \in \varphi$ that is violated by the input $(x, y)$.

Communication lower bounds for search problems have applications in many areas of complexity theory. We consider two areas that are the most relevant and explain the applicability of communication lower bounds.

**Proof complexity.** This area of complexity theory studies how hard it is to prove that a given formula $\varphi$ is unsatisfiable; in other words, what is the length of the shortest proof in a certain proof system. Lower bounds for the proof systems often correspond to lower bounds on a run-time of SAT-solvers, and there are intricate connections to other areas of complexity theory, such as, for example, circuit complexity.

There is a general framework for obtaining lower bounds on the length of the shortest proofs via communication. Suppose that, for an unsatisfiable CNF formula $\varphi$, we divide the variables into two disjoint groups $X$ and $Y$ in an arbitrary way. For a fixed proof system $\mathfrak{C}$ we can try to



ISSN 1433-8092

transform efficient proof of $\varphi$ into an efficient communication protocol for $\mathrm{Search}_\varphi$. A lower bound on the communication complexity of $\mathrm{Search}_\varphi$ then implies a lower bound on the length of a proof of $\varphi$ in $\mathfrak{C}$.

This framework seems to originate from [LNNW95]. Following this reduction, lower bounds for many different proof systems were obtained, for example: tree-like Cutting Planes [IPU94, HN12, dRNV16, BW25], tree-like Threshold proof system [BPS07], tree-like Res($\oplus$) [IS20], etc. [HN12, GP18a]. Depending on the communication model, even dag-like proofs can be analyzed via this framework [Kra97, Pud97, HP17, FPPR22, GGKS20, Sok24].

The lower bounds that can be achieved via this technique depend on the power of the communication model: the more powerful model we consider, the bigger class of proof systems we get the lower bound for. The choice of the formula $\varphi$ is important here as well, in a sense that we need to be able to show the lower bound on the communication complexity of $\mathrm{Search}_\varphi$. Typically, $\varphi$ is artificially built for this purpose. In this paper, we show a communication lower bound for the natural class of formulas (without usage of ad hoc constructions) that is a candidate for being hard for all propositional proof systems.

**Circuit complexity.** Natural embedding of $\mathrm{Search}_\varphi$ into a monotone Karchmer–Wigderson relation [KW90, Raz90] gives us the opportunity to use it for proving lower bounds for the monotone models of computation. From communication lower bounds, strong results are known for monotone formulas [RPRC16], monotone circuits [GGKS20, LMM+22], monotone span programs [RPRC16, PR18], etc. Communication is also the main instrument for showing separation between those models [PR18, GKRS19], and trade-off results [dRFJ+25, GMRS25]. These type of results are based on ad hoc constructions of the formulas $\varphi$. Namely, $\varphi$ is designed in order to able to show communication lower bound.

## 1.1 Random CNF

To be more precise we start with the definition of random CNF formulas.

**Definition 2.** Let $\mathfrak{F}(m, n, \Delta)$ denote the distribution of random $\Delta$-CNF on $n$ variables obtained by sampling $m$ clauses (out of the $\binom{n}{\Delta}2^\Delta$ possible clauses) uniformly at random with repetitions.

The famous result of Chvátal–Szemerédi says that if we pick a formula from this distribution with the proper parameters, the resulting formula will be unsatisfiable with high probability.

**Theorem 3** (Chvátal–Szemerédi, [CS88])**.** *For any $\Delta \geq 3$ whp $\boldsymbol{\varphi} \sim \mathfrak{F}(m, n, \Delta)$ is unsatisfiable if $m \geq \ln 2 \cdot 2^\Delta n$.*

These types of distributions appear not only in most of the areas in computer science, but in general mathematics and physics as well [MPZ02]. An interesting application is due to Feige [Fei02], who conjectured the following statement: no polynomial time algorithm may *prove* whp the unsatisfiability of a random $O(1)$-CNF formula with arbitrary large constant clause density. Assuming Feige's conjecture, it is known that some problems are hard to approximate: vertex covering, PAC learning DNFs [DSS16], etc.

As a candidate to be hard to refute in all proof systems, random CNFs are actively studied and lower bounds are known for many different proof systems [Gri01, BKPS02, AR03, Ale11, SS22]. Recent developments in this direction utilize the connection between proof complexity of $\varphi$ and

communication complexity of $\text{Search}_\varphi$. In particular, lower bounds for the Cutting Planes proofs of random $O(\log n)$-CNF [HP17, FPPR22, Sok24] follow this strategy. However, these results only consider lower bounds on deterministic *dag-like communication complexity* of $\text{Search}_\varphi$ based on random $O(\log n)$-CNF.

In this paper, we analyse the randomized tree-like communication of this problem that is incomparable with deterministic dag-like communication. This is a natural problem in a natural model, which also provides a way to explore how techniques used for structured formulas might extend to more typical instances like random CNFs. The main result is the following.

**Theorem 4.** *Let $c > 0$ be a large enough constant, $n > 0, \Delta \geq c \log n, m = O(n2^{\Delta})$. If $\boldsymbol{\varphi} \sim \mathfrak{F}(m, n, \Delta)$ and $\boldsymbol{X}, \boldsymbol{Y} \subseteq [n]$ is a partition of variables that is taken uniformly at random, then whp over choice of $\boldsymbol{\varphi}$ and partition $\boldsymbol{X}, \boldsymbol{Y}$ the randomized communication complexity of $\text{Search}_{\boldsymbol{\varphi}}$ is $\Omega(n)$.*

## 1.2 Prior Results and Technique

For several types of formulas $\varphi$, the randomized communication complexity of $\text{Search}_\varphi$ is well-studied. The approach for proving such bounds is the reduction of *Unique Disjointness* function to $\text{Search}_\varphi$. The main success in this direction is the reduction based on *critical block sensitivity* [HN12, GP18b], we also include some earlier results, though there is some difference in the technique [BPS07]. More precisely, for this technique one should assume that $\varphi = \psi \circ g$ (we take some formula $\psi$ and, in place of each variable, we substitute a carefully chosen gadget $g$ with fresh variables). Assuming that $\text{Search}_\psi$ has critical block sensitivity $m$ (that is a generalization of the block sensitivity measure), it is possible to reduce instances of unique disjointness of size $\text{poly}(m)$ to $\text{Search}_\varphi$.

The general framework for working with such formulas of $\psi \circ g$ is called *lifting*, and the idea is to "lift" the hardness of $\psi$ with respect to another complexity measure to communication complexity via gadget. Lifting can be based on the other complexity measures as well. For example, it can also be implemented for randomized decision tree complexity instead of critical block sensitivity [GJPW18]; however, this method requires the lower bound on the randomized decision tree complexity, which might be non-trivial, especially in case of $\text{Search}_\varphi$ problem. Such lower bound is known for Tseitin formulas [GJW18], together with [GJPW18] it yields the lower bound for randomized communication complexity of Search for Tseitin formulas lifted by Inner Product.

The notable exception here is the lower bound on Search problem for *Binary Pigeonhole Principle* (BPHP) [IR21]. These formulas are not lifted, however the proof is also the reduction of Unique Disjointness to the Search problem. This reduction based on the inner symmetry of BPHP.

A different kind of proof of a $\text{Search}_\varphi$ lower bound was given by Yang and Zhang [YZ24] (based on [WYZ23, YZ24]), who prove a lower bound for a weak version of BPHP. In contrast to the previous works this one is not a reduction from Unique Disjointness. Instead, they directly apply the structure versus randomness framework from the lifting literature [GLM+16, GJPW18] to the potential protocol that computes $\text{Search}_\varphi$.

Our proof of Theorem 4 combines the approach of [GLM+16, GJPW18, YZ24] with the analysis of expander graphs via closure argument that was developed for proof complexity purposes in [AR03, ABSRW04]. However, we use the iterative construction of the closure from [Sok20]. In part, this is also inspired by [GNRS24].

More precisely, the proof of our result is based on the following steps.

1. Following [HP17, FPPR22, Sok24] we divide variables between Alice and Bob uniformly at random.

2. Following the line of work on lifting of randomized decision trees [GLM+16, GJPW18, YZ24, GGJL25] we show that every communication protocol can be converted into a more structured one, a so-called subcube-like protocol. In such a communication protocol, each rectangle is a product of two sets with some bits fixed and the remaining pseudorandom.

3. Due to the nature of our random CNFs, the invariant that all clauses contain pseudorandom variables is not strong enough on its own. Search problem still might become trivial early on in communication protocol; for example, if the contradiction could be narrowed down to a small set of clauses. To avoid this problem, we use the closure trick [AR03, ABSRW04, Sok20, GNRS24], that allows us to maintain expansion property on the pseudorandom part of the graph.

4. Following [GGJL25], we show that the number of fixed bits in each rectangle is at most $O(d/\varepsilon)$ if we allow error $\varepsilon$, where $d$ is the communication complexity of the original protocol.

In addition, we show the better error bound dependency on the protocol depth $d$ than in [GGJL25]. We give a more refined analysis of the conversion to the subcube-like protocols. More precisely, we show that the number of fixed bits in each rectangle is $O(d)$ even when we allow for the $\exp(-d)$ error.

## 2 Notation and Tools

We denote the standard binary entropy function by $\mathrm{H}(p) \coloneqq p\log(1/p) + (1-p)\log(1/(1-p))$.

**Definition 5.** A bipartite graph $G = (L, R, E)$ is called an $(r, \Delta, \alpha\Delta)$-expander, if all vertices in $L$ have degree at most $\Delta$ and for any set $S \subseteq L$ such that $|S| \leq r$ it holds that $|N(S)| \geq \alpha\Delta|S|$, where $N_G(S)$ denotes the set of neighbours of $S$ in $G$ (we omit the subscript if the graph is clear from the context).

With a CNF formula $\varphi$ over $n$ variables and with $m$ clauses we associate a graph $G_\varphi \coloneqq ([m], [n], E)$ in a natural way: $(i, j) \in E$ iff the $i$-th clause contains the $j$-th variable. The following Lemma gives us some useful properties of underlying graphs of random CNFs. It follows from a standard computation, which was featured, for example, in [Sok24, Lemma A.2].

**Lemma 6.** *Let $n > 0$, $\eta > 0$ be an arbitrary constant, $\Delta = c\log n$, for a large enough constant $c$ depends on $\eta$, $m = O(n2^\Delta)$. Let $G \coloneqq ([m], [n], E)$ be a bipartite graph, such that each $i \in [m]$ choses $\Delta$ neighbours uniformly at random over $\binom{n}{\Delta}$ possibilities. Then $G$ is an $(r, \Delta, (1-\eta)\Delta)$-expander for $r = \Omega(n/\Delta)$.*

Instead of working directly with randomized communication, we use the equivalent characterization through distributional communication complexity. That is, we prove a lower bound against deterministic protocols that achieve error $\varepsilon$ with respect to a certain distribution on inputs (here we use uniform distribution), and a lower bound against randomized protocols that achieve error $\varepsilon$ follows. Below, "communication protocol" refers to a deterministic communication protocol.

# 3 Refuting Bipartite CNFs

In this section we mainly prove a special "bipartite" case of Theorem 4. We show in Section 3.1 that it actually implies the general case.

**Theorem 7.** *Let $\alpha > 0$ be an absolute constant. Let $G_1 \coloneqq ([m], [n], E_1), G_2 \coloneqq ([m], [n], E_2)$ be two $(r, \Delta, \alpha\Delta)$-expanders, and $X, Y \coloneqq \{0,1\}^n$. For each $i \in [m]$, let $C_i$ be a disjunction of variables in $\{x_j \mid j \in N_{G_1}(i)\} \cup \{y_j \mid j \in N_{G_2}(i)\}$ with arbitrary signs. Then for every communication protocol $\Pi\colon \{0,1\}^n \times \{0,1\}^n \to [m]$ of depth $d$ at most $O(\Delta r)$:*

$$\Pr_{\substack{(\boldsymbol{x},\boldsymbol{y})\sim X\times Y \\ \boldsymbol{i}\sim\Pi(\boldsymbol{x},\boldsymbol{y})}}[C_{\boldsymbol{i}}(\boldsymbol{x},\boldsymbol{y}) = 0] \le d \cdot 2^{-\Omega(\Delta)} + \exp(-d).$$

This section is organized as follows. In Section 3.1, we derive Theorem 4 from Theorem 7. In Section 3.2, we formally define subcube-like protocols and provide necessary tools. In Section 3.3, we give a more refined analysis of the conversion from general protocols to subcube-like ones in [GGJL25]. In Section 3.4, we show the hardness of $\mathrm{Search}_\varphi$ against subcube-like protocols when the underlying graphs are good expanders. Finally, in Section 3.5, we put everything together and derive Theorem 7.

## 3.1 Deriving Theorem 4 from Theorem 7

The main part of the argument that reduces the general case to the bipartite is a clean-up lemma essentially saying that incurring a small error we can treat the general case as bipartite. Similar arguments have been made in [HP17, FPPR22, Sok24].

Let $\varphi = \bigwedge_{i\in[n]} C_i$ be a $\Delta$-CNF with the set of variables $[n]$. Let $A \sqcup B = [n]$ be a partition of the variables. Let $G_A \coloneqq ([m], A, E_A)$ and $G_B \coloneqq ([m], B, E_B)$ be the graphs with edges connecting a clause with all variables from one of the sets mentioned in it. Let $\textsc{Error}_A \subseteq [m]$ and $\textsc{Error}_B \subseteq [m]$ be the sets of clauses with degree exceeding $(1-\delta)\Delta$ in $G_A$ and $G_B$ respectively. It means that clauses from $[m] \setminus (\textsc{Error}_A \cup \textsc{Error}_B)$ have at least $\delta\Delta$ variables from $A$ and $B$. We then say that $(A, B)$ is *$\delta$-good partition for $\varphi$* if

1. $\Pr_{\boldsymbol{x}\sim\{0,1\}^A}[\forall i \in \textsc{Error}_A \text{ we have } C_i(\boldsymbol{x}, \cdot) \equiv 1] \ge 1 - 2^{-\Omega(\Delta)}$.
2. $\Pr_{\boldsymbol{y}\sim\{0,1\}^B}[\forall i \in \textsc{Error}_B \text{ we have } C_i(\cdot, \boldsymbol{y}) \equiv 1] \ge 1 - 2^{-\Omega(\Delta)}$.
3. $G_A - \textsc{Error}_A - \textsc{Error}_B$ and $G_B - \textsc{Error}_A - \textsc{Error}_B$ are $(r, \Delta, \delta\Delta/2)$-expanders, where $r = \Omega(n/\Delta)$.

In this definition we assume that $\delta$ is an absolute constant and hidden constants depend on it.

**Lemma 8.** *Let $\boldsymbol{\varphi} \sim \mathfrak{F}(m, n, \Delta)$ with $\Delta = c \log n$ and $m = \alpha 2^{\Delta} n$, where $c, \alpha > 0$ are constants, and $c \ge 40$. Let $\boldsymbol{X}, \boldsymbol{Y}$ be a uniformly random partition of $[n]$. Then whp $(\boldsymbol{X}, \boldsymbol{Y})$ is a $\delta$-good partition for $\boldsymbol{\varphi}$ for any $\delta \le 1/10$.*

We defer the proof of this lemma to Appendix A.

*Proof of Theorem 4.* Applying Lemma 8, we get that the variable partition $\boldsymbol{X} \sqcup \boldsymbol{Y} = [n]$ is 1/10-good wrt $\boldsymbol{\varphi}$. Let $G_1 \coloneqq G_{\boldsymbol{X}} - \textsc{Error}_{\boldsymbol{X}} - \textsc{Error}_{\boldsymbol{Y}}$, $G_2 \coloneqq G_{\boldsymbol{Y}} - \textsc{Error}_{\boldsymbol{X}} - \textsc{Error}_{\boldsymbol{Y}}$. Note that the left parts of these graphs have equal size. We can add dummy variables to the right parts of these graphs to make them equal for the simplicity of notation.

By Lemma 8, the probability over $(\boldsymbol{x}, \boldsymbol{y}) \sim \{0,1\}^{\boldsymbol{X}} \times \{0,1\}^{\boldsymbol{Y}}$ for the $\textsc{Error}_{\boldsymbol{X}}$ or $\textsc{Error}_{\boldsymbol{Y}}$ to not be immediately satisfied is $2^{-\Omega(\Delta)}$. This means that if we consider a protocol for the problem $\mathrm{Search}_{\boldsymbol{\varphi}}$ for the variable partition $\boldsymbol{X} \sqcup \boldsymbol{Y}$ with the probability of success $\varepsilon$, we can reinterpret it as a protocol for $G_1$ and $G_2$ with the probability of success at least $\varepsilon - 2^{-\Omega(\Delta)}$.

We apply Theorem 7 with $\alpha \coloneqq \frac{1}{20}$. Since $r\Delta$ can be as large as $\Omega(n)$ by Lemma 6, we can pick the constants depending on $\alpha$ such that the probability from Theorem 7 is less than $\frac{1}{100}$. Then the probability of success for the problem from Theorem 4 is less than $\frac{1}{100} + 2^{-\Omega(\Delta)}$, and the theorem follows. □

### 3.2 Density Restoring Machinery

Every communication protocol $\Pi$ can be seen as a tree (not necessarily binary). Let $\mathcal{N}(\Pi)$ denote the set of all nodes in $\Pi$. Each node $v \in \mathcal{N}(\Pi)$ is associated with a rectangle, denoted $R_v = X_v \times Y_v$.

**Definition 9 (Min-entropy).** For a random variable $\boldsymbol{x}$, let $\mathbf{H}_\infty(\boldsymbol{x}) = \min_x \log \frac{1}{\Pr[\boldsymbol{x}=x]}$.

**Definition 10 (Spread variables).** Let $\boldsymbol{x} \in \{0,1\}^n$ be a random boolean vector. We say $\boldsymbol{x}$ is $\gamma$-*spread* if for every $I \subseteq [n]$ we have $\mathbf{H}_\infty(\boldsymbol{x}_I) \geq \gamma|I|$.

**Definition 11 (Structured variables).** Let $\boldsymbol{x} \in \{0,1\}^n$ be a random boolean vector and $I \subseteq [n]$. We say $\boldsymbol{x}$ is $(I,\gamma)$-*structured* if there exists some $a_I \in \{0,1\}^I$ such that

- $\Pr[\boldsymbol{x}_I = a_I] = 1$;
- $\boldsymbol{x}_{[m]\setminus I}$ is $\gamma$-spread.

**Definition 12 (Subcube-like rectangle).** A rectangle $R = X \times Y \subseteq \{0,1\}^n \times \{0,1\}^n$ is $\gamma$-subcube-like with respect to $(I,J)$ where $I, J \subseteq [n]$ if $\boldsymbol{x} \sim X$ is $(I,\gamma)$-structured and $\boldsymbol{y} \sim Y$ is $(J,\gamma)$-structured. In which case, we use $\mathsf{fix}(X) \coloneqq I$ and $\mathsf{fix}(Y) \coloneqq J$ to denote the fixed part of $X$ and $Y$ respectively.

**Definition 13 (Subcube-like protocols [GGJL25]).** A communication protocol $\Pi\colon \{0,1\}^n \times \{0,1\}^n \to S$ is $\gamma$-subcube-like if for every node $v \in \mathcal{N}(\Pi)$ in the protocol tree, $R_v$ is $\gamma$-subcube-like.

**Definition 14 (Codimension).** The codimension of a subcube-like rectangle $R = X \times Y$ is defined as the total number of fixed positions in $X$ and $Y$, denoted $\mathrm{codim}(R) \coloneqq |\mathsf{fix}(X)| + |\mathsf{fix}(Y)|$. The codimension of a subcube-like protocol $\Pi$ is the maximum codimension of subcube-like rectangles associated with any nodes in the protocol tree of $\Pi$, denoted $\mathrm{codim}(\Pi) \coloneqq \max_{v \in \mathcal{N}(\Pi)} \mathrm{codim}(R_v)$.

**Lemma 15** (Density Restoring Partition [GPW20])**.** *Let $\boldsymbol{x} \in \{0,1\}^n$ be a random boolean vector with support $X \subseteq \{0,1\}^n$ and $0 < \gamma < 1$ be a fixed parameter. There exists a partition*

$$X = X^1 \sqcup X^2 \sqcup X^3 \cdots \sqcup X^r$$

*such that for each $j \in [r]$, $\boldsymbol{x} \mid \boldsymbol{x} \in X^j$ is $(I^j,\gamma)$-structured with respect to some $I^j \subseteq [n]$.*

*Moreover, if we denote $p^{\geq j} \coloneqq \Pr\big[\boldsymbol{x} \in X^j \sqcup \cdots \sqcup X^\ell\big]$, then it holds that:*

$$\boldsymbol{H}_\infty(\boldsymbol{x}_{[m]\setminus I^j} \mid \boldsymbol{x} \in X^j) \geq \boldsymbol{H}_\infty(\boldsymbol{x}) - \gamma|I^j| - \log\big(1/p^{\geq j}\big).$$

### 3.3 Subcube-like protocols from general protocols

Göös et al. [GGJL25] show how to convert an arbitrary communication protocol into a subcube-like one. Specifically, they prove the following.

**Lemma 16** ([GGJL25])**.** *Let $\Pi$ be a communication protocol of depth $d$ and $\varepsilon > 0$. There exists a subcube-like protocol $\tilde{\Pi}$ of codimension $\operatorname{codim}(\tilde{\Pi}) = O(d/\varepsilon)$ such that*

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[\Pi(\boldsymbol{x},\boldsymbol{y}) \neq \tilde{\Pi}(\boldsymbol{x},\boldsymbol{y})] \leq \varepsilon.$$

Their bound is tight in the constant-error regime. However, it degenerates when $\varepsilon = O(d/n)$.

In this subsection, we give a more refined analysis of the reduction in [GGJL25], which makes the bound applicable in the inverse polynomial error regime (when $d = \Omega(\log n)$). We remark that such an analysis has been implicitly provided in [GPW20].

**Lemma 17.** *Let $\Pi$ be a communication protocol of depth $d$. There exists a $\gamma$-subcube-like protocol $\tilde{\Pi}$ of codimension $\operatorname{codim}(\tilde{\Pi}) = \frac{7}{1-\gamma} \cdot d$ such that*

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[\Pi(\boldsymbol{x},\boldsymbol{y}) \neq \tilde{\Pi}(\boldsymbol{x},\boldsymbol{y})] \leq \exp(-d).$$

We include a simplified version of the algorithm for such conversion from [GGJL25] for completeness. This algorithm simulates a subcube-like protocol $\Pi'$ on an input $(x, y)$, given a general protocol $\Pi$.

**Algorithm 2 (simplified) conversion from [GGJL25]**

$v \leftarrow$ root of $\Pi$, $X \times Y = \{0,1\}^n \times \{0,1\}^n$, $I \leftarrow \varnothing$.
**while** $v$ is not a leaf **do**
  $v_0, v_1 \leftarrow$ children of $v$
  **Suppose** Alice **sends** a bit at $v$ (otherwise swap $X$ and $Y$, $I$ and $J$)
  **Let** $X = X^0 \sqcup X^1$ be the partition according to the bit Alice sends
  **Let** $X^b = \bigsqcup_i X^{b,i}$ be the density-restoring partition (with parameter $\gamma$ and sets $I^i$, respectively), where $x \in X^b$.
  $X \leftarrow X^{b,i}, I \leftarrow I \cup I^i$ where $x \in X^i$
  Alice **sends** $(b, C(i))$ to Bob (here $C(i)$ is any encoding of $i$)
  $v \leftarrow v^b$
**end while**
**Output** the label $\Pi(v)$

*Proof.* Let $\Pi'$ be as given by Algorithm 2 in [GGJL25]. More precisely, let the protocol tree of $\Pi'$ consist all the possible configurations at the end of each iteration plus the initial one as the root (so the tree is not necessarily binary). Observe that $\Pi'$ has depth $d$, though may have much larger communication complexity.

For any $x, y \in \{0,1\}^n$, we have $\Pi(x,y) = \Pi'(x,y)$. It suffices to show $\Pr_{\boldsymbol{x},\boldsymbol{y}}[\operatorname{codim}(R(\boldsymbol{x},\boldsymbol{y})) > \frac{7}{1-\gamma} \cdot d] \leq \exp(-d)$, where $R(x,y)$ is the unique rectangle associated with the leaves of $\Pi$ that contains $(x,y)$. The desired $\tilde{\Pi}$ can then be obtained by shaving all the nodes in the protocol tree of $\Pi'$ associated with a rectangle of codimension greater than $\frac{7}{1-\gamma} \cdot d$.

For each node $v \in \mathcal{N}(\Pi')$, define the entropy deficiency of $v$ as

$$\mathbf{D}_\infty(v) \coloneqq \mathbf{D}_\infty(X_v) + \mathbf{D}_\infty(Y_v),$$

where

$$\mathbf{D}_\infty(X_v) \coloneqq n - |\mathsf{fix}(X_v)| - \mathbf{H}_\infty(\boldsymbol{x}_{[n]\setminus \mathsf{fix}(X_v)}), \qquad \boldsymbol{x} \sim X_v$$

and $\mathbf{D}_\infty(Y_v)$ is defined analogously.

Now consider running $\Pi'$ on $x, y$, and let $v_0, \ldots, v_d \in \mathcal{N}(\Pi')$ denote all the nodes on the execution path. Fix any $k \in [d]$ and let us simply use $u$ and $v$ to denote $v_{k-1}$ and $v_k$. Suppose without loss of generality that it is Alice who sends a bit to Bob in the $k$-th iteration. Recall that in each iteration Alice first partitions $X_u = X_u^0 \sqcup X_u^1$ according to the bit she sends. Then she performs the density-restoring partition with parameter $\gamma$ on $X_u^b = X_u^{b,1} \sqcup \ldots \sqcup X_u^{b,r}$ where $x \in X_u^b$. Finally, she determines the unique $X_u^{b,i}$ that contains $x$. Then for the next configuration, $X_v = X_u^{b,i}$. Let us define

$$\begin{aligned}
q_u^b &\coloneqq \Pr\Big[\boldsymbol{x} \in X_u^b \mid \boldsymbol{x} \in X_u\Big], \\
p_u^{b,\geq j} &\coloneqq \Pr\left[\boldsymbol{x} \in \bigcup_{k \geq j} X_u^{b,k} \middle| \boldsymbol{x} \in X_u^b\right] \quad \forall j \in [r], \\
h_k(x,y) &\coloneqq \log\Big(1/q_u^b\Big) + \log\Big(1/p_u^{b,\geq i}\Big), \\
n_k(x,y) &\coloneqq |\mathsf{fix}(X_v) \setminus \mathsf{fix}(X_u)|.
\end{aligned}$$

We have the following simple fact.

**Fact 18.** *$\boldsymbol{D}_\infty(v) \leq \boldsymbol{D}_\infty(u) - (1-\gamma) n_k(x,y) + h_k(x,y)$.*

*Proof.* For proof see appendix D. □

Together with the nonnegativity of $\mathbf{D}_\infty$, we can bound the codimension of $R(x,y)$ by $h(x,y) \coloneqq \sum_{k=1}^d h_k(x,y)$ up to a multiplicative factor.

**Claim 19.** *For every $x, y \in \{0,1\}^n$, $\operatorname{codim}(R(x,y)) \leq \frac{1}{1-\gamma} \cdot h(x,y)$.*

*Proof.* Consider the path in the tree leading to the leaf $R(x,y)$, this path being of length $d$. Summing up the inequalities from Fact 18 along that path, we get:

$$\mathbf{D}_\infty(v_d) - \mathbf{D}_\infty(v_0) \leq -(1-\gamma) \sum_{j=1}^d n_j(x,y) + \sum_{j=1}^d h_j(x,y)$$

Since $\mathbf{D}_\infty(v_d)$ is non-negative and $\mathbf{D}_\infty(v_0) = 0$, it follows that:

$$\operatorname{codim}(R(x,y)) = \sum_{j=1}^d n_j(x,y) \leq \frac{1}{1-\gamma} \sum_{j=1}^d h_j(x,y) = \frac{1}{1-\gamma} \cdot h(x,y).$$ □

We also observe that $h_k(\boldsymbol{x}, \boldsymbol{y})$ has an exponential tail for each $k \in [d]$, even conditioned on any node $v$ of depth $k-1$ being reached.

**Claim 20.** *For every node $v \in \mathcal{N}(\Pi')$ of depth $0 \le k < d$ and threshold $\gamma \ge 0$,*

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[h_{k+1}(\boldsymbol{x},\boldsymbol{y}) \ge 1+\gamma \mid \boldsymbol{v}_k = v] \le 2^{-\gamma}.$$

*Proof.* Let $\boldsymbol{b}, \boldsymbol{i}$ be as defined in the $(k+1)$-th iteration of Algorithm 2 given $\boldsymbol{x}, \boldsymbol{y}$. We have

$$\begin{aligned}
&\Pr_{\boldsymbol{x},\boldsymbol{y}}[h_{k+1}(\boldsymbol{x},\boldsymbol{y}) \ge 1+\gamma \mid \boldsymbol{v}_k = v] \\
=& \sum_{b\in\{0,1\}} \Pr[\boldsymbol{b} = b \mid \boldsymbol{v}_k = v] \cdot \Pr\left[\log\left(1/q_v^b\right) + \log\left(1/p_v^{b,\ge \boldsymbol{i}}\right) \ge 1+\gamma \mid \boldsymbol{b} = b, \boldsymbol{v}_k = v\right] \\
=& \sum_{b\in\{0,1\}} q_v^b \cdot \Pr\left[q_v^b \cdot p^{b,\ge \boldsymbol{i}} \le 2^{-\gamma-1} \middle| \boldsymbol{b} = b, \boldsymbol{v}_k = v\right] \\
\le& \sum_{b\in\{0,1\}} q_v^b \cdot \min\left\{1, 2^{-\gamma-1} \cdot q_v^b\right\} \\
\le& 2^{-\gamma},
\end{aligned}$$

where in the second last inequality, we use the property that $\Pr\left[p_v^{b,\ge \boldsymbol{i}} \le t \mid \boldsymbol{b} = b, \boldsymbol{v}_k = v\right] \le t$ for all $t \in [0,1]$. □

Finally, we need the following adaptive version of Bernstein's inequality, whose proof can be found in Appendix C.

**Lemma 21.** *Let $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in \mathbb{R}$ be a random sequence of reals and $\zeta > 0$ be some fixed parameter. If for any $1 \le k \le n$ and $a_1, \dots, a_{k-1} \in \mathbb{R}$ such that $\Pr[\boldsymbol{a}_1 = a_1, \dots, \boldsymbol{a}_{k-1} = a_{k-1}] > 0$,*

$$\Pr[\boldsymbol{a}_k \ge x \mid \boldsymbol{a}_1 = a_1, \dots, \boldsymbol{a}_{k-1} = a_{k-1}] \le \exp(-\zeta x),$$

*then*

$$\Pr\left[\sum_{i=1}^n \boldsymbol{a}_i \ge \frac{4}{\zeta} n\right] \le \exp(-n).$$

We are now ready to bound the codimension of $R(\boldsymbol{x},\boldsymbol{y})$. Let $(\boldsymbol{a}_k \coloneqq h_k(\boldsymbol{x},\boldsymbol{y}) - 1)_{k\in[d]} \in \mathbb{R}^d$ be a random sequence of reals. By Claim 20, $\boldsymbol{a}$ satisfies the condition in Lemma 21 with $\zeta = \ln 2$. Therefore,

$$\Pr[h(\boldsymbol{x},\boldsymbol{y}) \ge 7d] = \Pr\left[\sum_{i=1}^d \boldsymbol{a}_i \ge 6d\right] \le \exp(-d).$$

Together with Claim 19, we conclude that

$$\Pr\left[\operatorname{codim}(R(\boldsymbol{x},\boldsymbol{y})) \ge \frac{7}{1-\gamma} \cdot d\right] \le \Pr[h(\boldsymbol{x},\boldsymbol{y}) \ge 7d] \le \exp(-d).$$ □

### 3.4 Lower bound against subcube-like protocols

The following lemma is implicit in [GNRS24], we include its proof in Appendix B for completeness. In fact

**Lemma 22.** *Let $0 < \beta < \alpha < 1$ and let $\Pi\colon \{0,1\}^n \times \{0,1\}^n \to S$ be a subcube-like protocol of codimension $d := \mathrm{codim}(\Pi)$ where $d \le (\alpha-\beta)^2 r\Delta/4$, and $G_1 = ([m],[n],E_1), G_2 = ([m],[n],E_2)$ be two $(r,\Delta,\alpha\Delta)$-expanders. Then there exist families $\{\mathsf{Cl}^X(v)\}_{v\in\mathcal{N}(\Pi)}, \{\mathsf{Cl}^Y(v)\}_{v\in\mathcal{N}(\Pi)}$ of subsets of $[m]$ such that the following conditions hold:*

1. *For every non-root $v \in \mathcal{N}(\Pi)$, let $u$ denote $v$'s parent. Then $\mathsf{Cl}^X(u) \subseteq \mathsf{Cl}^X(v)$ and $\mathsf{Cl}^Y(u) \subseteq \mathsf{Cl}^Y(v)$.*
2. *For every $v \in \mathcal{N}(\Pi)$, $G_1 - \mathsf{Cl}^X(v) - N(\mathsf{Cl}^X(v)) - \mathsf{fix}(X_v)$ and $G_2 - \mathsf{Cl}^Y(v) - N(\mathsf{Cl}^Y(v)) - \mathsf{fix}(Y_v)$ are both $(r,\Delta,\beta\Delta)$-expanders.*
3. *For every $v \in \mathcal{N}(\Pi)$, $|\mathsf{Cl}^X(v)|, |\mathsf{Cl}^Y(v)| \le \frac{1}{\alpha-\beta}\cdot\frac{d}{\Delta}$.*

**Lemma 23.** *As in Theorem 7 let $G_1 := ([m],[n],E_1), G_2 := ([m],[n],E_2)$ be two $(r,\Delta,\alpha\Delta)$-expanders, and $X, Y := \{0,1\}^n$. For each $i \in [m]$, let $C_i$ be a disjunction of variables in $\{x_j \mid j \in N_{G_1}(i)\} \cup \{y_j \mid j \in N_{G_2}(i)\}$ with arbitrary signs. Let $\Pi\colon \{0,1\}^n \times \{0,1\}^n \to [m]$ be a $\gamma$-subcube-like communication protocol of $d := \mathrm{codim}(\Pi)$. If $d \le \alpha^2 r\Delta/4$, then*

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[C_{\boldsymbol{i}}(\boldsymbol{x},\boldsymbol{y}) = 0 \mid \boldsymbol{i} = \Pi(\boldsymbol{x},\boldsymbol{y})] \le O(2^{-\gamma\alpha\Delta/2}\cdot d).$$

*Proof.* We rephrase the success probability of $\Pi$ as follows: Sample a random leaf $\boldsymbol{\ell}$ of $\Pi$ with probability $|R_{\boldsymbol{\ell}}|/2^{2n}$. Then

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[C_{\boldsymbol{i}}(\boldsymbol{x},\boldsymbol{y}) = 0 \mid \boldsymbol{i} = \Pi(\boldsymbol{x},\boldsymbol{y})] = \mathbb{E}_{\boldsymbol{\ell}}\left[\Pr_{(\boldsymbol{x},\boldsymbol{y})\sim R_{\boldsymbol{\ell}}}[C_{\Pi(\boldsymbol{\ell})}(\boldsymbol{x},\boldsymbol{y}) = 0]\right]. \tag{1}$$

Let $\{\mathsf{Cl}^X(v)\}_{v\in\mathcal{N}(\Pi)}, \{\mathsf{Cl}^Y(v)\}_{v\in\mathcal{N}(\Pi)}$ be given by Lemma 22 with respect to $\Pi, G_1, G_2$ and $\beta = \alpha/2$. For each node $v \in \mathcal{N}(\Pi)$, define $J_v := \mathsf{Cl}^X(v) \cup \mathsf{Cl}^Y(v)$. We first observe that for each leaf $\ell$, $\Pi$ has low success probability on $R_\ell$ if $\Pi(\ell) \notin J_\ell$.

**Claim 24.** *Let $\ell$ be any leaf in the protocol tree of $\Pi$. Suppose that $i \notin J_\ell$. Then*

$$\Pr_{(\boldsymbol{x},\boldsymbol{y})\sim R_\ell}[C_i(\boldsymbol{x},\boldsymbol{y}) = 0] \le 2^{-\gamma\alpha\Delta/2}.$$

*Proof.* By the definition of $J_\ell$, we have $i \notin \mathsf{Cl}^X(\ell)$. Let $A \subseteq [n] \setminus (\mathsf{fix}(X) \cup N(\mathsf{Cl}^X(\ell)))$ be the set of neighbors of $i$ in $G_1 - \mathsf{Cl}^X(\ell) - N(\mathsf{Cl}^X(\ell)) - \mathsf{fix}(X_\ell)$, by the expansion we get $|A| \ge \alpha\Delta/2$. Since $X_\ell \times Y_\ell$ is $\gamma$-subcube-like we have that $\boldsymbol{x}_{[n]\setminus\mathsf{fix}(X_\ell)}$ is $\gamma$-spread. In particular, $\mathbf{H}_\infty(\boldsymbol{x}_A) \ge \gamma|A| \ge \gamma\alpha\Delta/2$. Let $\tau \in \{0,1\}^A$ be the unique assignment that violates all literals of $C_i$ in $A$. The min-entropy bound above then implies $\Pr[C_i(\boldsymbol{x},\boldsymbol{y}) = 0] \le \Pr[\boldsymbol{x}_A = \tau] \le 2^{-\gamma\alpha\Delta/2}$. □

On the other hand, unfortunately, it is possible that

$$p_i(\ell) := \Pr_{(\boldsymbol{x},\boldsymbol{y})\sim R_\ell}[C_i(\boldsymbol{x},\boldsymbol{y}) = 0]$$

is close to 1 for some $i \in J_\ell$. Nevertheless, we can show that this can happen only for a small fraction of leaves.

**Claim 25.** *Let $\boldsymbol{\ell}$ be a random leaf sampled as stated above. Then*

$$\mathbb{E}_{\boldsymbol{\ell}}\left[\sum_{i\in J_{\boldsymbol{\ell}}} p_i(\boldsymbol{\ell})\right] \le 2^{-\gamma\alpha\Delta/2}\cdot d.$$

*Proof.* First, we can write

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\ell}}\left[\sum_{i\in J_{\boldsymbol{\ell}}} p_i(\boldsymbol{\ell})\right] &= \sum_{i\in[m]} \mathbb{E}_{\boldsymbol{\ell}}[\mathbf{1}_{i\in J_{\boldsymbol{\ell}}}\cdot p_i(\boldsymbol{\ell})] \\
\text{(where } \ell_{\boldsymbol{x},\boldsymbol{y}} \text{ is the leaf containing } (\boldsymbol{x},\boldsymbol{y})) &= \sum_{i\in[m]} \Pr_{\boldsymbol{x},\boldsymbol{y}}[i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}} \wedge C_i(\boldsymbol{x},\boldsymbol{y})=0] \\
&= \sum_{i\in[m]} \Pr_{\boldsymbol{x},\boldsymbol{y}}[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}]\cdot \Pr_{\boldsymbol{x},\boldsymbol{y}}[i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}] \\
&\le \max_{i\in[m]} \Pr_{\boldsymbol{x},\boldsymbol{y}}[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}]\cdot \sum_{i\in[m]} \Pr_{\boldsymbol{x},\boldsymbol{y}}[i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}] \\
&= \max_{i\in[m]} \Pr_{\boldsymbol{x},\boldsymbol{y}}[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}]\cdot \mathbb{E}_{\boldsymbol{\ell}}[|J_{\boldsymbol{\ell}}|].
\end{aligned}$$

Observe that $|J_\ell| \le d$ for every leaf $\ell$, it suffices to show

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}] \le 2^{-\gamma\alpha\Delta/2}.$$

for every $i\in[m]$. Now let us fix an arbitrary $i\in[m]$. The event "$i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}$" can be reinterpreted as follows: with

$$V_i \coloneqq \{v\in\mathcal{N}(\Pi) \mid i\in J_v \text{ and the parent of } v \text{ does not satisfy that}\}$$

we have that $i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}$ if and only if $(\boldsymbol{x},\boldsymbol{y})\in\bigsqcup_{v\in V_i} R_v$ (the rectangles form a partition since the nodes in $V_i$ are maximally close to the root). Then

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}] \le \sum_{v\in V_i} \Pr_{\boldsymbol{x},\boldsymbol{y}}[(\boldsymbol{x},\boldsymbol{y})\in R_v \mid i\in J_{\ell_{\boldsymbol{x},\boldsymbol{y}}}]\cdot \Pr_{\boldsymbol{x},\boldsymbol{y}}[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid (\boldsymbol{x},\boldsymbol{y})\in R_v].$$

Since the right-hand side is a convex combination of $\Pr[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid (\boldsymbol{x},\boldsymbol{y})\in R_v]$ for $v\in V_i$, it suffices to bound the maximum of these probabilities.

The crucial observation to conclude the proof is that $i\notin \mathsf{Cl}^X(v)$ if Bob spoke in the parent node of $v$ and $i\notin \mathsf{Cl}^Y(v)$ if Alice spoke in that node. In any case, an argument similar to that in Claim 24 applies and we have $\Pr[C_i(\boldsymbol{x},\boldsymbol{y})=0 \mid (\boldsymbol{x},\boldsymbol{y})\in R_v] \le 2^{-\gamma\alpha\Delta/2}$, which concludes the proof. □

Now we are ready to show the desired bound. Combining the above two claims, we have

$$\begin{aligned}
(1) &= \Pr[\Pi(\boldsymbol{\ell})\notin J_{\boldsymbol{\ell}}]\cdot \mathbb{E}_{\boldsymbol{\ell}}[\Pr_{(\boldsymbol{x},\boldsymbol{y})\sim R_{\boldsymbol{\ell}}}[C_{\Pi(\boldsymbol{\ell})}(\boldsymbol{x},\boldsymbol{y})=0] \mid \Pi(\boldsymbol{\ell})\notin J_{\boldsymbol{\ell}}] \\
&\quad + \Pr[\Pi(\boldsymbol{\ell})\in J_{\boldsymbol{\ell}}]\cdot \mathbb{E}_{\boldsymbol{\ell}}[\Pr_{(\boldsymbol{x},\boldsymbol{y})\sim R_{\boldsymbol{\ell}}}[C_{\Pi(\boldsymbol{\ell})}(\boldsymbol{x},\boldsymbol{y})=0] \mid \Pi(\boldsymbol{\ell})\in J_{\boldsymbol{\ell}}] \\
&\le 2^{-\gamma\alpha\Delta/2} + \mathbb{E}_{\boldsymbol{\ell}}[\sum_{i\in J_{\boldsymbol{\ell}}} p_i(\boldsymbol{\ell})] \\
&= O(d/2^{\gamma\alpha\Delta/2}).
\end{aligned}$$

□

### 3.5 Proof of Theorem 7

We first restate the theorem for convenience.

**Theorem 7.** *Let $\alpha > 0$ be an absolute constant. Let $G_1 \coloneqq ([m], [n], E_1), G_2 \coloneqq ([m], [n], E_2)$ be two $(r, \Delta, \alpha\Delta)$-expanders, and $X, Y \coloneqq \{0,1\}^n$. For each $i \in [m]$, let $C_i$ be a disjunction of variables in $\{x_j \mid j \in N_{G_1}(i)\} \cup \{y_j \mid j \in N_{G_2}(i)\}$ with arbitrary signs. Then for every communication protocol $\Pi\colon \{0,1\}^n \times \{0,1\}^n \to [m]$ of depth $d$ at most $O(\Delta r)$:*

$$\Pr_{\substack{(\boldsymbol{x},\boldsymbol{y})\sim X\times Y \\ \boldsymbol{i}\sim\Pi(\boldsymbol{x},\boldsymbol{y})}}[C_{\boldsymbol{i}}(\boldsymbol{x},\boldsymbol{y}) = 0] \le d \cdot 2^{-\Omega(\Delta)} + \exp(-d).$$

*Proof.* Let $\tilde{\Pi}$ be a subcube-like protocol given by Lemma 17 with respect to $\Pi$ and $\gamma = \alpha$. Then $\operatorname{codim}(\Pi') \le \frac{7d}{1-\alpha}$. Moreover,

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[\Pi(\boldsymbol{x},\boldsymbol{y}) \neq \tilde{\Pi}(\boldsymbol{x},\boldsymbol{y})] \le \exp(-d).$$

We can then apply Lemma 23 and conclude that

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}[C_{\boldsymbol{i}}(\boldsymbol{x},\boldsymbol{y}) = 0 \mid \boldsymbol{i} = \Pi(\boldsymbol{x},\boldsymbol{y})] \le \Pr_{\boldsymbol{x},\boldsymbol{y}}[C_{\boldsymbol{i}}(\boldsymbol{x},\boldsymbol{y}) = 0 \mid \boldsymbol{i} = \tilde{\Pi}(\boldsymbol{x},\boldsymbol{y})] + \exp(-d) = d\cdot 2^{-\Omega(\Delta)} + \exp(-d). \square$$

## References


[ABSRW04] Michael Alekhnovich, Eli Ben-Sasson, Alexander A. Razborov, and Avi Wigderson. Pseudorandom generators in propositional proof complexity. *SIAM J. Comput.*, 34(1):67–88, 2004. doi:10.1137/S0097539701389944.

[Ale11] Michael Alekhnovich. Lower bounds for k-dnf resolution on random 3-cnfs. *Comput. Complex.*, 20(4):597–614, 2011. doi:10.1007/s00037-011-0026-0.

[AR03] Michael Alekhnovich and Alexander A. Razborov. Lower bounds for polynomial calculus: Non-binomial case. *Proceedings of the Steklov Institute of Mathematics*, 242:18–35, 2003. Available at http://people.cs.uchicago.edu/~razborov/files/misha.pdf. Preliminary version in *FOCS '01*.

[BKPS02] Paul Beame, Richard M. Karp, Toniann Pitassi, and Michael E. Saks. The efficiency of resolution and davis–putnam procedures. *SIAM J. Comput.*, 31(4):1048–1075, 2002. doi:10.1137/S0097539700369156.

[BPS07] Paul Beame, Toniann Pitassi, and Nathan Segerlind. Lower bounds for lov[a-acute]sz–schrijver systems and beyond follow from multiparty communication complexity. *SIAM J. Comput.*, 37(3):845–869, 2007. doi:10.1137/060654645.

[BW25] Paul Beame and Michael Whitmeyer. Multiparty Communication Complexity of Collision-Finding and Cutting Planes Proofs of Concise Pigeonhole Principles. In Keren Censor-Hillel, Fabrizio Grandoni, Joël Ouaknine, and Gabriele Puppis, editors, *52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025)*, volume 334 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages

21:1–21:20, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ICALP.2025.21, doi:10.4230/LIPIcs.ICALP.2025.21.

[CS88] Vašek Chvátal and Endre Szemerédi. Many hard examples for resolution. *J. ACM*, 35(4):759–768, October 1988. URL: http://doi.acm.org/10.1145/48014.48016, doi:10.1145/48014.48016.

[dRFJ+25] Susanna F. de Rezende, Noah Fleming, Duri Andrea Janett, Jakob Nordström, and Shuo Pang. Truly supercritical trade-offs for resolution, cutting planes, monotone circuits, and weisfeiler–leman. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1371–1382, New York, NY, USA, 2025. Association for Computing Machinery. doi:10.1145/3717823.3718271.

[dRNV16] Susanna F. de Rezende, Jakob Nordström, and Marc Vinyals. How limited interaction hinders real communication (and what it means for proof and circuit complexity). In Irit Dinur, editor, *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 295–304. IEEE Computer Society, 2016. doi:10.1109/FOCS.2016.40.

[DSS16] Amit Daniely and Shai Shalev-Shwartz. Complexity theoretic limitations on learning dnf's. In Vitaly Feldman, Alexander Rakhlin, and Ohad Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 815–830, Columbia University, New York, New York, USA, 23–26 Jun 2016. PMLR. URL: https://proceedings.mlr.press/v49/daniely16.html.

[Fei02] Uriel Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity, Montréal, Québec, Canada, May 21-24, 2002*, page 5. IEEE Computer Society, 2002. URL: http://doi.ieeecomputersociety.org/10.1109/CCC.2002.10006, doi:10.1109/CCC.2002.10006.

[FPPR22] Noah Fleming, Denis Pankratov, Toniann Pitassi, and Robert Robere. Random $\Theta(\log n)$-CNFs are Hard for Cutting Planes. *J. ACM*, 69(3):19:1–19:32, 2022. doi:10.1145/3486680.

[GGJL25] Mika Göös, Tom Gur, Siddhartha Jain, and Jiawei Li. Quantum communication advantage in tfnp. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1465–1475, New York, NY, USA, 2025. Association for Computing Machinery. doi:10.1145/3717823.3718155.

[GGKS20] Ankit Garg, Mika Göös, Pritish Kamath, and Dmitry Sokolov. Monotone circuit lower bounds from resolution. *Theory Comput.*, 16:1–30, 2020. URL: https://doi.org/10.4086/toc.2020.v016a013, doi:10.4086/TOC.2020.V016A013.

[GJPW18] Mika Göös, T. S. Jayram, Toniann Pitassi, and Thomas Watson. Randomized communication versus partition number. *ACM Trans. Comput. Theory*, 10(1):4:1–4:20, 2018. doi:10.1145/3170711.

[GJW18] Mika Göös, Rahul Jain, and Thomas Watson. Extension complexity of independent set polytopes. *SIAM J. Comput.*, 47(1):241–269, 2018. doi:10.1137/16M109884X.

[GKRS19] Mika Göös, Pritish Kamath, Robert Robere, and Dmitry Sokolov. Adventures in monotone complexity and TFNP. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 38:1–38:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. URL: https://doi.org/10.4230/LIPIcs.ITCS.2019.38, doi:10.4230/LIPICS.ITCS.2019.38.

[GLM+16] Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. *SIAM J. Comput.*, 45(5):1835–1869, 2016. doi:10.1137/15M103145X.

[GMRS25] Mika Göös, Gilbert Maystre, Kilian Risse, and Dmitry Sokolov. Supercritical tradeoffs for monotone circuits. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1359–1370, New York, NY, USA, 2025. Association for Computing Machinery. doi:10.1145/3717823.3718229.

[GNRS24] Mika Göös, Ilan Newman, Artur Riazanov, and Dmitry Sokolov. Hardness condensation by restriction. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 2016–2027. ACM, 2024. doi:10.1145/3618260.3649711.

[GP18a] Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. *SIAM J. Comput.*, 47(5):1778–1806, 2018. doi:10.1137/16M1082007.

[GP18b] Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. *SIAM Journal on Computing*, 47(5):1778–1806, 2018. doi:10.1137/16M1082007.

[GPW20] Mika Göös, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for BPP. *SIAM Journal on Computing*, 49(4), 2020. doi:10.1137/17M115339X.

[Gri01] Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1):613–622, 2001. URL: http://www.sciencedirect.com/science/article/pii/S0304397500001572, doi:10.1016/S0304-3975(00)00157-2.

[HN12] Trinh Huynh and Jakob Nordström. On the virtue of succinct proofs: amplifying communication complexity hardness to time-space trade-offs in proof complexity. In Howard J. Karloff and Toniann Pitassi, editors, *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 233–248. ACM, 2012. doi:10.1145/2213977.2214000.

[HP17] Pavel Hrubes and Pavel Pudlák. Random formulas, monotone circuits, and interpolation. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 121–131. IEEE Computer Society, 2017. doi:10.1109/FOCS.2017.20.

[IPU94] Russell Impagliazzo, Toniann Pitassi, and Alasdair Urquhart. Upper and lower bounds for tree-like cutting planes proofs. In *Proceedings of the Ninth Annual Symposium on Logic in Computer Science (LICS '94), Paris, France, July 4-7, 1994*, pages 220–228. IEEE Computer Society, 1994. doi:10.1109/LICS.1994.316069.

[IR21] Dmitry Itsykson and Artur Riazanov. Proof complexity of natural formulas via communication arguments. In Valentine Kabanets, editor, *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, volume 200 of *LIPIcs*, pages 3:1–3:34. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. URL: https://doi.org/10.4230/LIPIcs.CCC.2021.3, doi:10.4230/LIPICS.CCC.2021.3.

[IS20] Dmitry Itsykson and Dmitry Sokolov. Resolution over linear equations modulo two. *Ann. Pure Appl. Log.*, 171(1), 2020. URL: https://doi.org/10.1016/j.apal.2019.102722, doi:10.1016/J.APAL.2019.102722.

[Kra97] Jan Krajícek. Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. *J. Symb. Log.*, 62(2):457–486, 1997. doi:10.2307/2275541.

[KW90] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM J. Discret. Math.*, 3(2):255–265, 1990. doi:10.1137/0403021.

[LMM+22] Shachar Lovett, Raghu Meka, Ian Mertz, Toniann Pitassi, and Jiapeng Zhang. Lifting with sunflowers. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPIcs*, pages 104:1–104:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. URL: https://doi.org/10.4230/LIPIcs.ITCS.2022.104, doi:10.4230/LIPICS.ITCS.2022.104.

[LNNW95] László Lovász, Moni Naor, Ilan Newman, and Avi Wigderson. Search problems in the decision tree model. *SIAM J. Discret. Math.*, 8(1):119–132, 1995. doi:10.1137/S0895480192233867.

[MPZ02] Marc Mezard, Giorgio Parisi, and Riccardo Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science (New York, N.Y.)*, 297:812–815, 09 2002. doi:10.1126/science.1073287.

[PR18] Toniann Pitassi and Robert Robere. Lifting nullstellensatz to monotone span programs over any field. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1207–1219. ACM, 2018. doi:10.1145/3188745.3188914.

[Pud97] Pavel Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *J. Symb. Log.*, 62(3):981–998, 1997. doi:10.2307/2275583.

[Raz90] Alexander A. Razborov. Applications of matrix methods to the theory of lower bounds in computational complexity. *Comb.*, 10(1):81–93, 1990. doi:10.1007/BF02122698.

[RPRC16] Robert Robere, Toniann Pitassi, Benjamin Rossman, and Stephen A. Cook. Exponential lower bounds for monotone span programs. In Irit Dinur, editor, *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 406–415. IEEE Computer Society, 2016. doi:10.1109/FOCS.2016.51.

[Sok20] Dmitry Sokolov. (semi)algebraic proofs over ±1 variables. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 78–90. ACM, 2020. doi:10.1145/3357713.3384288.

[Sok24] Dmitry Sokolov. Random $(\log n)$-CNF Are Hard for Cutting Planes (Again). In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 2008–2015. ACM, 2024. doi:10.1145/3618260.3649636.

[SS22] Anastasia Sofronova and Dmitry Sokolov. A lower bound for $k$-dnf resolution on random CNF formulas via expansion. *Electron. Colloquium Comput. Complex.*, TR22-054, 2022. URL: https://eccc.weizmann.ac.il/report/2022/054.

[WYZ23] Shuo Wang, Guangxu Yang, and Jiapeng Zhang. Communication Complexity of Set-Intersection Problems and Its Applications. Technical report, ECCC, 2023. URL: https://eccc.weizmann.ac.il/report/2023/164.

[YZ24] Guangxu Yang and Jiapeng Zhang. Communication lower bounds for collision problems via density increment arguments. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 630–639, New York, NY, USA, 2024. Association for Computing Machinery. doi:10.1145/3618260.3649607.

# A Proof of Lemma 8

We start by proving Item 1 (Item 2 is analogous). Let $N^{\boldsymbol{X}}(i)$ for $i \in [m]$ be the set of neighbors of $i$ in $\boldsymbol{X}$ and $N^{\boldsymbol{Y}}(i)$ — in $\boldsymbol{Y}$. Then $\textsc{Error}_{\boldsymbol{X}} = \{i \in [m] \mid |N^{\boldsymbol{X}}(i)| > (1-\delta)\Delta\}$. We then write

$$\begin{aligned}
\mathbb{E}[|\textsc{Error}_{\boldsymbol{X}}|] &= \sum_{i\in[m]} \Pr\left[|N^{\boldsymbol{X}}(i)| > (1-\delta)\Delta\right] \\
&= \sum_{i\in[m]} \sum_{S\subseteq N(i)\colon |S|\geq(1-\delta)\Delta} \Pr[\boldsymbol{X} \cup N(i) = S] \\
&= m \sum_{j\leq\delta\Delta} \binom{m}{m-j} 2^{-\Delta} \\
&\leq m 2^{-(1-\mathrm{H}(\delta))\Delta} \\
&= \alpha n \cdot 2^{\mathrm{H}(\delta)\Delta} \\
&= \alpha n^{1+c\,\mathrm{H}(\delta)}
\end{aligned}$$

On the other hand for every $i \in \textsc{Error}_{\boldsymbol{X}}$ we have

$$\Pr_{\boldsymbol{x}\sim\{0,1\}^{\boldsymbol{X}}}[C_i(\boldsymbol{x},\cdot) \not\equiv 1] = 2^{-|N^{\boldsymbol{X}}(i)|} \le 2^{-(1-\delta)\Delta} = n^{-c(1-\delta)}.$$

Then by a union bound we get

$$\Pr_{\boldsymbol{x}\sim\{0,1\}^{\boldsymbol{X}}}[\exists i \in \textsc{Error}_{\boldsymbol{X}} \colon C_i(\boldsymbol{x},\cdot) \not\equiv 1] \le |\textsc{Error}_{\boldsymbol{X}}| \cdot n^{-c(1-\delta)}.$$

Then by Markov's inequality applied to $|\textsc{Error}_{\boldsymbol{X}}|$ with probability $1-\varepsilon$ over $\boldsymbol{X}$ we have

$$\begin{aligned}\Pr_{\boldsymbol{x}\sim\{0,1\}^{\boldsymbol{X}}}[\exists i \in \textsc{Error}_{\boldsymbol{X}} \colon C_i(\boldsymbol{x},\cdot) \not\equiv 1] &\le 1/\varepsilon \cdot \mathbb{E}[|\textsc{Error}_{\boldsymbol{X}}|] \cdot n^{-c(1-\delta)} \\ &= \alpha/\varepsilon \cdot n^{1+c\,\mathrm{H}(\delta)-c(1-\delta)} \\ &= \alpha/\varepsilon \cdot n^{1-c(1-\delta-\mathrm{H}(\delta))}\end{aligned}$$

Now it remains to prove Item 3. First, let $\boldsymbol{G} \coloneqq ([m],[n],E_{\boldsymbol{X}} \sqcup E_{\boldsymbol{Y}})$ be the union of $G_{\boldsymbol{X}}$ and $G_{\boldsymbol{Y}}$. By Lemma 6 whp over $\boldsymbol{\varphi}$ the graph $\boldsymbol{G}$ is an $(r,\Delta,(1-\eta)\Delta)$-expander for any $\eta$ and $r = \Omega_\eta\left(\frac{n}{\Delta}\right)$. Now it is sufficient to show $G_{\boldsymbol{X}} - \textsc{Error}_{\boldsymbol{Y}}$ is an $(r,\Delta,(\delta-2\eta)\Delta)$-expander whp, $G_{\boldsymbol{Y}} - \textsc{Error}_{\boldsymbol{X}}$ is distributed identically to $G_{\boldsymbol{X}} - \textsc{Error}_{\boldsymbol{Y}}$ and removing additional nodes from the left-hand side does not reduce expansion.

We in fact show that conditioned on the fact that $\boldsymbol{G}$ is an $(r,\Delta,(1-\eta)\Delta)$-expander, $G_{\boldsymbol{X}} - \textsc{Error}_{\boldsymbol{Y}}$ is $(r,\Delta,(\delta-2\eta)\Delta)$-expander with probability 1. Consider an arbitrary subset $U \subseteq [m]$ of size at most $r$. For every such subset we need to have $|N(U) \setminus \boldsymbol{Y} \setminus N(\textsc{Error}_{\boldsymbol{Y}})| \ge (\delta - 2\eta)\Delta|U \setminus \textsc{Error}_{\boldsymbol{Y}}|$. Here and below $N(S) = N_{\boldsymbol{G}}(S)$. Consider the set $\partial U \coloneqq \{v \in N(U) \mid v \text{ is connected with a single node in } U\}$. Then $|\partial U| \ge (1-2\eta)|U|\Delta$: indeed the number of edges incident to $U$ can be estimated in two ways:

$$\Delta|U| = |E \cap (U \times [n])| \ge |\partial U| + 2(|N(U)| - |\partial U|) \ge 2(1-\eta)\Delta|U| - |\partial U|.$$

Then we can partition the set $N(U)$ into sets $N_i$ for $i \in U$ where $N_i \subseteq N(i)$ and $|N_i| \ge (1-2\eta)\Delta$: find a node $i \in U$ such that $|\partial U \cap N(i)| \ge (1-2\eta)\Delta$, let $N_i \coloneqq \partial U \cap N(i)$ and continue the process for $U \setminus \{i\}$, the reason the resulting sets form a partition is that $N_i \cap N(U \setminus \{i\}) = \emptyset$ by the definition of $\partial U$.

For every $Y \subseteq [n]$ if $|N(i) \setminus Y| \ge \delta\Delta$, then $|N_i \setminus Y| \ge |N(i) \setminus Y| - |N(i) \setminus N_i| \ge (\delta - 2\eta)\Delta$. It follows that after removing $\textsc{Error}_{\boldsymbol{Y}}$ (vertices for which $|N(i) \setminus \boldsymbol{Y}| < \delta\Delta$), every vertex $i$ in the set $U \setminus \textsc{Error}_{\boldsymbol{Y}}$ in $G_{\boldsymbol{X}}$ has at least $(\delta-2\eta)\Delta$ neighbours in $N_i \setminus \boldsymbol{Y}$. As $N_i$ is a partition, it follows that $|N(U) \setminus \textsc{Error}_{\boldsymbol{Y}} \setminus N(\textsc{Error}_{\boldsymbol{Y}})| \ge (\delta-2\eta)\Delta|U \setminus \textsc{Error}_{\boldsymbol{Y}}|$.

Finally, choosing $\eta = \delta/4$ completes the proof. By Lemma 6 this particular choice only affects the hidden constant in $r = \Omega(n/\Delta)$.

# B Proof of Lemma 22

Since $\mathsf{Cl}^X$ and $\mathsf{Cl}^Y$ are independent of each other, we just focus on constructing $\mathsf{Cl}^X$. It suffices to prove the following lemma:

**Lemma 26.** *Let $G = ([m], [n], E)$ be an $(r, \Delta, \alpha\Delta)$-expander. Let $\mathcal{T}$ be a tree with nodes labeled with subsets of $[n]$, where $S_v \subseteq [n]$ denotes the label of $v$ such that*

- *For the root of $\mathcal{T}$, the node $r$ we have $S_r = \emptyset$.*
- *If $u$ is a parent of $v$, then $S_u \subseteq S_v$.*
- *For every $u$ we have $|S_u| \leq d \leq (\alpha - \beta)^2 r\Delta/4$.*

*Then there for every node $u$ to $\mathcal{T}$ there exists a set $T_u \subseteq [m]$ such that*

*(a) The graph $G_u \coloneqq G - T_u - S_u - N(T_u)$ is an $(r, \Delta, \beta\Delta)$-expander.*

*(b) If $u$ is a parent of $v$, then $T_u \subseteq T_v$.*

*(c) $|T_u| \leq \frac{1}{\alpha-\beta} d/\Delta$.*

To finish the proof of Lemma 22 given Lemma 26 we just let $\mathcal{T}$ be the tree of the protocol and $S_u$ be $\mathsf{fix}(X_u)$, then take $\mathsf{Cl}^X(u) \coloneqq T_u$.

We now proceed to prove Lemma 26. Wlog we may assume that if $u$ is a parent of $v$ we have $|S_v \setminus S_u| \leq 1$ (just by replacing a single edge in $\mathcal{T}$ by a chain of edges).

We construct the sets $T_u$ inductively starting from the root $r$ where $T_r = \emptyset$. Suppose $u$ is a parent of node $v$ and we have constructed $T_u$. If $S_u = S_v$, we just let $T_v \coloneqq T_u$, so assume that $S_v \setminus S_u = \{i\}$. Let $G'_u \coloneqq G_u - i$. Let us find the largest set $B_v \subseteq [m] \setminus T_u$ such that $|B_v| \leq r$ and $|N_{G'_u}(B_v)| \leq \beta\Delta|B_v|$ and let $T_v \coloneqq T_u \cup B_v$. Then $G_v = G'_u - T_u - N_{G'_u}(T_u)$. It is clear that $T$ satisfies Item (b).

**Proof of Item (c)** We show by induction on the depth $\ell$ of a node $u$ that $|T_u| \leq \frac{1}{\alpha-\beta}\ell/\Delta$. The base case is satisfied since for the root $r$ the set $T_r$ is empty. Now let $u$ be a node at depth $\ell$ and $v$ be its child at depth $\ell + 1$. We have that $|T_u| \leq \frac{1}{\alpha-\beta}\ell/\Delta$, we need to prove that $|T_v| = |T_u \sqcup B_v| \leq \frac{1}{\alpha-\beta}(\ell + 1)/\Delta$.

On the one hand $N_{G'_u}(B_v) = N_G(B_v) \setminus (N_G(T_u) \cup S_v)$. On the other hand, $|N_{G'_u}(B_v)| \leq \beta\Delta|B_v|$. By the expansion of $G$ we have $|N_G(B_v)| \geq \alpha\Delta|B_v|$. Hence $|N_G(T_u) \cup S_v| \geq (\alpha - \beta)\Delta|B_v|$. By the assumption on the tree $|S_v| = \ell + 1$, and by induction hypothesis $|T_u| \leq \frac{1}{\alpha-\beta}\ell/\Delta$, so $|N_G(T_u))| \leq \frac{1}{\alpha-\beta}\ell$.

Combining the two inequalities, we get $\frac{1}{\alpha-\beta}\ell + (\ell + 1) \geq (\alpha - \beta)\Delta|B_v|$.

From that, we get $|B_v| \leq 2 \cdot \frac{1}{(\alpha-\beta)^2\Delta} \cdot (\ell + 1) \leq r/2$, where the last inequality follows from the assumptions on $\ell$. Then we get that $|T_v| \leq |T_u| + |B_v| \leq r$. Now we can use expansion of $G$ to bound $|N_G(T_v)| \geq \alpha\Delta|T_v|$. On the other hand, let $r = w_0, w_1, \ldots, w_\ell = u, w_{\ell+1} = v$ be the path in $\mathcal{T}$ from the root to $v$. We then have

$$N_G(T_v) \subseteq \bigcup_{i=0}^{\ell} N_{G'_{w_i}}(B_{w_{i+1}}) \cup S_v.$$

By the choice of sets $B$ we get $|N_G(T_v)| \leq \beta\Delta|T_v| + |S_v|$. Combining the two bounds we get $|T_v| \leq \frac{1}{\alpha-\beta}|S_v|/\Delta$, which concludes the proof.

**Proof of Item (a)** Pick the node $v$ at depth $\ell+1$ such that $G_v$ is not an $(r, \Delta, \beta\Delta)$-expander, and $v$ is the closest to the root among such nodes. In particular, for its parent $u$ the graph $G_u$ is $(r, \Delta, \beta\Delta)$-expander. Then there exists a set $T$ of size at most $r$ such that $N_{G_v}(T) < \beta\Delta|T|$. By expansion of $G$ we get $|N_G(T)| \geq \alpha\Delta|T|$. Then, since $N_{G_v}(T) = N_G(T) \setminus (N_G(T_v) \cup S_v)$ we have

$$\frac{1}{2}(\alpha-\beta)\Delta r \geq \frac{2}{\alpha-\beta}\ell \geq \frac{1}{\alpha-\beta}\ell + (\ell+1) \geq |N_G(T_v) \cup S_v| \geq (\alpha-\beta)\Delta|T|.$$

The left-hand side follows from Item (c) and the right-hand side follows from the analysis above. Then $|T| \leq r/2$. Since by the proof of Item (c) we have that $|B_v| \leq r/2$, we get $|T \cup B_v| \leq r$, yet $|N_{G'_u}(T \cup B_v)| < \beta\Delta|B_v| + \beta\Delta|T| \leq \beta\Delta|B_v \sqcup T|$, contradicting the choice of $B_v$.

# C Proof of Lemma 21

**Lemma 21.** *Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}$ be a random sequence of reals and $\zeta > 0$ be some fixed parameter. If for any $1 \leq k \leq n$ and $a_1, \ldots, a_{k-1} \in \mathbb{R}$ such that $\Pr[\boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{k-1} = a_{k-1}] > 0$,*

$$\Pr[\boldsymbol{a}_k \geq x \mid \boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{k-1} = a_{k-1}] \leq \exp(-\zeta x),$$

*then*

$$\Pr\left[\sum_{i=1}^{n} \boldsymbol{a}_i \geq \frac{4}{\zeta} n\right] \leq \exp(-n).$$

*Proof.* Let $\lambda \in (0, \zeta)$ be some parameter that will be determined later. First, observe that for any $1 \leq k \leq n$ and $a_1, \ldots, a_{k-1} \in \mathbb{R}$ such that $\Pr[\boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{k-1} = a_{k-1}] > 0$,

$$\mathbb{E}[\exp(\lambda \boldsymbol{a}_k) \mid \boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{k-1} = a_{k-1}] \leq \zeta \int_0^\infty \exp(\lambda x) \cdot \exp(-\zeta x) \cdot \mathrm{d}x = \zeta/(\zeta-\lambda). \tag{2}$$

Next, we prove by induction on $k$ from $n$ to 1 that

$$\mathbb{E}\left[\exp\left(\lambda \sum_{i=k}^{n} \boldsymbol{a}_i\right) \middle| \boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{k-1} = a_{k-1}\right] \leq \left(\frac{\zeta}{\zeta-\lambda}\right)^{n-k+1}. \tag{3}$$

The base case $k = n$ is exactly (2). Now assume that (3) holds for all $k \geq m+1$. Then

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\lambda \sum_{i=m}^{n} \boldsymbol{a}_i\right) \middle| \boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{m-1} = a_{m-1}\right] \\
&\leq \sum_{a_m} \Pr[\boldsymbol{a}_m = a_m \mid \boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{m-1} = a_{m-1}] \cdot \exp(\lambda a_m) \cdot \\
&\qquad \cdot \mathbb{E}\left[\exp\left(\lambda \sum_{i=m+1}^{n} \boldsymbol{a}_i\right) \middle| \boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_m = a_m\right] \\
&\leq \left(\frac{\lambda}{\zeta-\lambda}\right)^{n-m} \cdot \mathbb{E}[\exp(\lambda \boldsymbol{a}_m) \mid \boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_{m-1} = a_{m-1}] \\
&= \left(\frac{\lambda}{\zeta-\lambda}\right)^{n-m+1}.
\end{aligned}$$

Finally, by setting $\lambda = \zeta/2$, we conclude that

$$\begin{aligned}
\Pr\left[\sum_{i=1}^{n} \boldsymbol{a}_i \geq \frac{4}{\zeta} n\right] &= \Pr\left[\exp\left(\lambda \sum_{i=1}^{n} \boldsymbol{a}_i\right) \geq \exp\left(\frac{4\lambda}{\zeta} n\right)\right] \\
&\leq \mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n} \boldsymbol{a}_i\right)\right] \cdot \exp(-2n) \\
&\leq \left(\frac{\zeta}{\zeta - \lambda}\right)^{n} \cdot \exp(-2n) \\
&\leq \exp(-n).
\end{aligned}$$

□

# D Proof of Fact 18

Follows from the computation:

$$\begin{aligned}
\mathbf{D}_{\infty}(v) - \mathbf{D}_{\infty}(u) &= -|\mathsf{fix}(X_v)| + |\mathsf{fix}(X_u)| + \mathbf{H}_{\infty}(X_v) - \mathbf{H}_{\infty}(X_u) \\
&= -n_k(x,y) + (\mathbf{H}_{\infty}(X_v) - \mathbf{H}_{\infty}(X_u^b)) + (\mathbf{H}_{\infty}(X_u^b) - \mathbf{H}_{\infty}(X_v)) \\
&\leq -n_k + \log\left(1/q_u^b\right) + \left(\gamma \cdot n_k(x,y) + \log\left(1/p_u^{b,\geq i}\right)\right) \qquad \text{(from Lemma 15)} \\
&= -(1-\gamma) n_k(x,y) + h_k(x,y).
\end{aligned}$$



ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il