Electronic Colloquium on Computational Complexity, Report No. 106 (2025)

Exponential Lower Bounds on the Size of ResLin Proofs of Nearly Quadratic Depth

Sreejata Kishor Bhattacharya * Arkadev Chattopadhyay †

July 30, 2025

Abstract

Itsykson and Sokolov [IS14] identified resolution over parities, denoted by Res(⊕), as a natural and simple fragment of $\mathrm{AC}^0[2]$-Frege for which no super-polynomial lower bounds on size of proofs are known. Building on a recent line of work, Efremenko and Itsykson [EI25] proved lower bounds of the form $\exp(N^{\Omega(1)})$, on the size of Res(⊕) proofs whose depth is upper bounded by $O(N \log N)$, where N is the number of variables of the unsatisfiable CNF formula. The hard formula they used was Tseitin on an appropriately expanding graph, lifted by a 2-stifling gadget. They posed the natural problem of proving super-polynomial lower bounds on the size of proofs that are $\Omega(N^{1+\epsilon})$ deep, for any constant $\epsilon > 0$.

We provide a significant improvement by proving a lower bound on size of the form $\exp(\tilde{\Omega}(N^{\epsilon}))$, as long as the depth of the Res(⊕) proofs are $O(N^{2-\epsilon})$, for every $\epsilon > 0$. Our hard formula is again Tseitin on an expander graph, albeit lifted with a different type of gadget. Our gadget needs to have small correlation with all parities.

An important ingredient in our work is to show that arbitrary distributions *lifted* with such gadgets fool *safe* affine spaces, an idea which originates in the earlier work of Bhattacharya, Chattopadhyay and Dvorak [BCD24].

1 Introduction

One of the simplest proof systems in propositional proof complexity is Resolution. Haken [Hak85] obtained the first super-polynomial lower bounds on the size of proofs in this system for a CNF encoding of the pigeon-hole-principle forty years ago. Since then it has been very well studied with many beautiful results (see for example [Urq87], [BW01], [ABRW04]). Yet, seemingly slight strengthenings of resolution seem to frustrate current techniques in obtaining non-trivial lower bounds. We will consider one such strengthening, that was introduced by Itsykson and Sokolov [IS14], about ten years ago. This system is called resolution over parities, abbreviated by Res(⊕) and denoted by Res(⊕). It augments resolution by allowing the prover to make $\mathbb{F}_2$-linear inferences, while working with $\mathbb{F}_2$-linear clauses. Proving superpolynomial lower bounds for Res(⊕) remains a challenge. It is easy to see that Res(⊕) is a subsystem of $AC^0[2]$-Frege. While we know strong lower bounds for AC^0-Frege (see for example [BIK+92]), obtaining super-polynomial lower bounds for $AC^0[2]$-Frege for any unsatisfiable formula in CNF would be a major breakthrough (see for example [MP97]). Thus, Res(⊕) is in some sense the weakest natural subfragment of $AC^0[2]$-Frege for which we don't currently have lower bounds.

*TIFR, Mumbai. Email: sreejata.bhattacharya@tifr.res.in. Supported by the Department of Atomic Energy, Govmt. of India, under project #RTI4001 and by a Google PhD Fellowship.

†TIFR, Mumbai. Email: arkadev.c@tifr.res.in. Supported by the Department of Atomic Energy, Govmt. of India, under project #RTI4001 and by a Google India Faculty Award.

ISSN 1433-8092

Proving lower bounds for Res(⊕) would be a stepping stone towards proving lower bounds for $AC^0[2]$-Frege.

Itsykson and Sokolov proved exponential lower bounds on the size of tree-like Res(⊕) proofs using customized arguments for some formulas. General techniques for tree-like Res(⊕) proofs were developed in the independent works of Beame and Koroth [BK23] and Chattopadhyay, Mande, Sanyal and Sherif [CMSS23] that lifted lower bounds on the height of ordinary tree-like resolution proofs of a formula to that of the size of tree-like Res(⊕) proofs of the same formula lifted with an appropriate gadget. A more recent line of work ([EGI24], [BCD24], [AI25], [EI25]) has focused on proving lower bounds against subsystems of Res(⊕) that are stronger than tree-like but weaker than general Res(⊕). Gryaznov, Pudlak and Talebanfard [GPT22] had proposed several notions of regular proofs for the Res(⊕) system as appropriate first target for proving lower bounds. Efremenko, Garlik and Itsykson [EGI24] established lower bounds against such a subsystem of Res(⊕) known as bottom-regular Res(⊕). Bhattacharya, Chattopadhyay and Dvorak [BCD24] exhibited a CNF which is easy for resolution but hard for bottom-regular Res(⊕) - thereby strictly separating unrestricted Res(⊕) from bottom-regular Res(⊕). Subsequently, Alekseev and Itsykson [AI25] significantly extended the reach of techniques by showing $\exp(N^{\Omega(1)})$ lower bounds against Res(⊕) refutations whose depth is restricted to be at most $O(N \log\log N)$, where N is the number of variables of the unsatisfiable CNF. This depth restriction was further improved to $O(N \log N)$ by Efremenko and Itsykson [EI25].

A natural way towards proving lower bounds for unrestricted Res(⊕) would be improving the depth restriction all the way to $N^{\omega(1)}$. However, the techniques of Efremenko and Itsykson [EI25] seem to get stuck at $O(N \log N)$. Efremenko and Itsykson [EI25] posed the natural open problem of proving superpolynomial lower bounds against Res(⊕) refutations whose depth is restricted to $O(N^{1+\epsilon})$ where $\epsilon > 0$ is some constant .

Our main result, stated below, achieves such a bound.

Theorem 1.1. *Let Φ be the Tseitin contradiction on a $(|V|, d, \lambda)$ expander with $0 < \lambda < 1$ a small enough constant and $|V|$ odd. Let $n = |V|d/2$ be the number of edges (which is also the number of variables in Φ). Let IP be the inner product gadget on $b = 500\log(n)$ bits. Let $\Psi = \Phi \odot IP$ be the lift of Φ by IP. Let $N = nb$ be the number of variables in Ψ. Then, any Res(⊕) refutation of Ψ of depth $\leq O(N^{2-\epsilon})$ requires size* $\exp(\tilde{\Omega}(N^{\epsilon}))$

This pushes the frontier of depth of proofs against which super-polynomial lower bounds on size for Res(⊕) can be obtained, from $O(N\log(N))$ to $\tilde{O}(N^2)$. Another way of interpreting our result is to say that any Res(⊕) proof of the hard formula Ψ of size $\exp(N^{o(1)})$ has to be almost N^2 deep, which is significantly super-critical.

Our work combines the approaches of Alekseev and Itsykson [AI25], Efremenko and Itsykson [EI25] and Bhattacharya, Chattopadhyay and Dvorak [BCD24] - along with a new equidistribution lemma for *safe* affine spaces.

1.1 Some Other Related Work

Our work makes use of the notion of amortized closure that was introduced by Efremenko and Itsykson [EI25]. Apart from improving the depth lower bounds of small size Res(⊕) proofs, [EI25] used this notion to give an alternative proof of a lifting theorem of Chattopadhyay and Dvorák [CD25] and their proof works for a broader class of gadgets. The lifting theorem is used in [CD25] to prove super-criticial tradeoffs between depth and size of tree-like Res(⊕) proofs.

Our work also crucially uses lifted distributions to boost the success probability of random walks with restarts. In particular, it uses an analytic property of the gadget to argue equiditribution of pre-images in a *safe* affine space in the lifted world of a point $z \in \{0,1\}^n$ in the unlifted world. Such equidistribution, albeit wrt rectangles, have been earlier implicitly proved (see for example [GLM+16, CFK+21]) as well as explicitly proved in [CDK+17]. The analytic property of the gadget used in these works was essentially small discrepancy wrt rectangles (or being a 2-source extractor), something that seems to be significantly stronger than what we need of the gadget in this work.

2 Organization

We have introduced preliminaries (basic notation, facts about closure, amortized closure etc in) Section 3. After that we have included a high level overview of our proof in Section 4. We describe the CNF in our final result in Section 6. We prove our final result (Theorem 1.1) in Section 9, which uses machinery developed in Sections 5, 7, 8. In the remaining part of this Section we describe what each of Sections 5, 7, 8 do. To gain a better understanding of the picture, the reader is advised to read the high level overview in Section 4 before continuing with this section.

In Section 5 we establish the conditional fooling lemma (Lemma 5.1) - which gives us the freedom to work with any lifted distribution. Along the way we shall prove an equidistribution lemma (Lemma 5.6) - this lemma will be used later in another part of the proof.

In Section 6 we describe our CNF. We also mention the properties of expanders that we shall use in our proof.

In Section 7 we formalize the idea of [AI25] used to establish lower-bounds for depth-restricted Res($\oplus$) refutations combined with our equidistribution lemma. To instantiate this approach, one has to specify a set of partial assignments $P \subseteq \{0, 1, *\}^n$ with certain properties. We mention those properties in Section 7, and show how this implies a lower-bound for depth-restricted Res($\oplus$).

In Section 8 we find a set of partial assignments with the aforementioned properties when the base CNF is the Tseitin contradiction over an expander graph. As mentioned in Section 4, given any partial assignment $\rho \in P \subseteq \{0, 1, *\}^n$, we have to construct a distribution $\mu = \mu_\rho$ on the unfixed coordinates satisfying certain properties. The last of these properties talks about the inability of low-depth parity decision trees to perform a certain task (on average) when the input is sampled from $G^{-1}(\mu)$. For ease of presentation, in Section 8 we first prove an unlifted analogue of that property: here, we establish the inability of low-depth ordinary decision trees to perform the analogous task in the unlifted world, when the input is chosen from μ. Then, we shall sketch how to modify the proof (in a white-box manner) to prove the original requirement of hardness against parity decision trees. The main technical tool required while modifying the proof to the lifted world is the equidistribution lemma (Lemma 5.6) established in Section 5. For completeness, we have included a self-contained proof of hardness against PDTs in Appendix B.

In Section 9, we put everything together to prove our final result (Theorem 1.1): existence of a CNF on N variables such that any Res($\oplus$) refutation of depth $N^{2-\epsilon}$ requires size $\exp(\tilde{\Omega}(N^\epsilon))$.

3 Preliminaries

3.1 General Notation

- For a probability distribution μ, when we sample a point x according to μ, we denote it by $x \leftarrow \mu$.
- When x is sampled according to uniform distribution over a set T, we denote it by $x \sim T$.

3.2 Resolution over parities

Definition 3.1. A linear clause ℓ_C is an expression of the form

$$\ell_C(x) = [\langle \ell_1, x\rangle = b_1] \vee [\langle \ell_2, x\rangle = b_2] \cdots \vee [\langle l_k, x\rangle = b_k]$$

Here $x, \ell_1, \cdots, \ell_k \in \mathbb{F}_2^n$. Note that the negation of ℓ_C, $\neg \ell_C$ is an affine space:

$$\neg \ell_C = \{x \in \mathbb{F}_2^n | \langle \ell_1, x\rangle = 1 - b_1, \cdots, \langle \ell_k, x\rangle = 1 - b_k\}$$

Also notice that every ordinary clause is also a linear clause.

Res($\oplus$) (defined in [IS14]) is a proof system where every proof line is a linear clause. The derivation rules are as follows:

1. **Weakening:** From ℓ_C, derive $\ell'_C(x) = \ell_C(x) \vee [\langle \ell, x\rangle = b]$
2. **Resolution:** From $\ell_C^{(1)}(x) = \ell_C(x) \vee [\langle \ell, x\rangle = b]$ and $\ell_C^{(2)}(x) = \ell_c(x) \vee [\langle \ell, x\rangle = 1 - b]$, derive $\ell_C(x)$

A Res($\oplus$) refutation of a CNF Φ starts with the axioms being the clauses of Φ (which, as noted above, are also linear clauses) and applies a sequence of derivation rules to obtain the empty linear clause $\emptyset$.

Affine DAGs

For an unsatisfiable CNF Φ define the search problem

$$\text{Search}(\Phi) = \{(x, C) | C \text{ is a clause of } \Phi, C(x) = 0\}.$$

Just as a resolution refutation of Φ can be viewed as a cube-DAG for solving Search(Φ), a Res($\oplus$) refutation can be viewed as an affine-DAG for solving Search(Φ).

Definition 3.2. An affine DAG for Search(Φ) is a DAG where there is a distinguished root r, each node v has an associated affine space A_v, and each node has outdegree either 2, 1, or 0. Each outdegree 0 node w is labelled with a clause of Φ, C_w. The following requirements are satisfied:

1. If v has two children v_1, v_2 then $A_v = A_{v_1} \cup A_{v_2}$.
2. If v has only one child w, then $A_v \subseteq A_w$.
3. If v has no children, then for any $x \in A_v$, $C_v(x) = 0$ where C_v is the clause labelled on v.
4. The affine space labelled on the root is the entire space $\mathbb{F}_2^n$.

A Res($\oplus$) refutation for Φ can be viewed as an affine DAG for Search(Φ) by viewing the sequence of derivations as a DAG: for each node, the associated affine space is the negation of the linear clause derived at that node. The leaves are the axioms - the clause labelled at each leaf is simply the corresponding axiom.

We classify nodes based on their outdegree as follows.

1. A node with no children is called a leaf.
2. A node with one child is called a *weakening node*. (Because in the Res($\oplus$) refutation this node was derived by weakening.)
3. Let v be a node with two children v_1, v_2. In this case it holds that $A_v = A_{v_1} \cup A_{v_2}$; $A_{v_1} = A_v \wedge [\langle \ell, x\rangle = b]$ and $A_{v_2} = A_v \wedge [\langle \ell, x\rangle = 1 - b]$ for some $\ell \in \mathbb{F}_2^n$. Such a node is called a *query node*; we say the affine DAG queries ℓ at node v. (In the Res($\oplus$) refutation, node v was obtained by resolving the linear form $\langle \ell, x\rangle$.)

Path of an input

Here we consider any affine DAG that arises from some Res($\oplus$) refutation. For any node v and any $x \in A_v$, we define the path of x starting from v as follows:

- Start with the currrent node v.
- If the current node is w and w has no children, terminate the path.
- If the current node is w and has two children w_1, w_2, we know that $A_w = A_{w_1} \cup A_{w_2}$. In this case it will hold that $A_{w_1} = A_w \cap \{\tilde{x} | \langle \ell, \tilde{x} \rangle = b\}$ and $A_{w_2} = A_w \cap \{\tilde{x} | \langle \ell, \tilde{x} \rangle = 1 - b\}$ for some $\ell \in \mathbb{F}_2$. If $\langle \ell, x \rangle = b$, the next node in the path is w_1. Otherwise, the next node in the path is w_2
- If the current node w has only one child w_1, the next node in the path is w_1.

The way the path is defined ensures that if the path of x visits the node w, $x \in A_w$. Consequently, for any x and v such that $x \in A_v$, the path of x starting from v visits a leaf whose clause is falsified by x.

In particular, for any x, if we follow the path traversed by x from the root, we end up at a clause falsified by x.

Definition 3.3. We define the length of a path to be the number of query nodes encountered on the path. (The weakening nodes do not contribute to the length.)

Definition 3.4. The depth of a node v is the largest length of a path from the root to v. The depth of the refutation is the depth of the deepest node.

3.3 Lifting CNFs

Definition 3.5. For a base CNF Φ on variables $\{z_1, z_2, \cdots, z_n\}$ and a gadget $g : \mathbb{F}_2^n \to \mathbb{F}_2$ we define the lifted CNF $\Phi \circ g$ as follows.

- The set of variables is $\{x_{i,j} | i \in [n], j \in [b]\}$
- For each clause C in Φ, we define the set of clauses $C \circ g$ as follows: let the variables involved in C be $\{x_i | i \in S\}$ and let $\alpha \in \mathbb{F}_2^S$ be the unique assignment to those variables that falsifies C. The set of clauses $C \circ g$ will involve variables from the set $\{x_{i,j} | i \in S, j \in [b]\}$. For every choice of $(a_i | i \in S)$ where $a_i \in g^{-1}(\alpha_i)$ we add the following clause to $C \circ g$:

$$\vee_{i \in S} \left[\vee_{j \in [b]} [x_{i,j} \neq \alpha_{i,j}] \right]$$

- The lifted CNF $\Phi \circ g$ is the conjunction of $C \circ g$ for every $C \in \Phi$.

The semantic interpretation of $\Phi \circ g$ is as follows:

$$\Phi \circ g(x) = \Phi(g(x_{1,1}, x_{1,2}, \cdots, x_{1,b}), g(x_{2,1}, x_{2,2} \cdots, x_{2,b}), \cdots, g(x_{n,1}, x_{n,2}, \cdots, x_{n,b}))$$

Thus if Φ is unsatisfiable, so is $\Phi \circ g$.

If the largest width of a clause in Φ is w and Φ has m clauses, the number of clauses in $\Phi \circ g$ will be at most mb^w. In particular, if $m \leq \text{poly}(n), b = O(\log(n))$ and $w = O(1)$ then the number of clauses of $\Phi \circ g$ is bounded by $\text{poly}(n)$.

3.4 Notations about lifted spaces

In this paper, we shall be working with a gadget $g : \mathbb{F}_2^b \to \mathbb{F}_2$. The base space will be $\mathbb{F}_2^n$. The lifted space will be $\mathbb{F}_2^N$ where $N = nb$. The coordinates of the lifted space are $\{(i, j) | i \in [n], j \in [b]\}$.

Definition 3.6. The set of coordinates $\{x_{i,j} | j \in [b]\}$ is called to be *the block of* i.

Definition 3.7. For any set $S \subseteq [n]$, the set $\text{VARS}(S)$ is defined as the following set of coordinates in the lifted space: $\text{VARS}(S) = \{x_{i,j} | i \in S, j \in [b]\}$.

Definition 3.8. For any point $x \in \mathbb{F}_2^{nb}$ and $i \in [n]$, define $x(i) = (x_{i,1}, x_{i,2}, \cdots, x_{i,b})$.

The gadget g naturally induces a function $g^n : \mathbb{F}_2^{nb} \to \mathbb{F}_2^n$ by independent applications of g on the n different blocks. We shall abbreviate g^n by G.

Definition 3.9. For any assignment $\beta \in \mathbb{F}_2^{\text{VARS}(S)}$ to the variables in blocks of S, we define the partial assignment $G(\beta) \in \{0, 1, *\}$ as follows:

$$G(\beta)_i = \begin{cases} * & \text{if } i \notin S \\ g(\beta(i)) & \text{otherwise} \end{cases}$$

Definition 3.10. For any partial assignment $\alpha \in \{0, 1, *\}^n$ define $G^{-1}(\alpha)$ as follows:

- Let $S = \text{supp}(\alpha)$. Then, $G^{-1}(\alpha) \subseteq \mathbb{F}_2^{\text{VARS}(S)}$:

$$G^{-1}(\alpha) = \{y | y \in \mathbb{F}_2^{\text{VARS}(S)}, g(y(i)) = \alpha_i \ \forall i \in S\}$$

Definition 3.11. For any distribution μ on $\mathbb{F}_2^n$ define the lifted distribution $G^{-1}(\mu)$ on $\mathbb{F}_2^{nb}$ as the outcome of the following sampling procedure:

1. Sample $z \leftarrow \mu$.
2. Sample x uniformly at random from $G^{-1}(z)$.

Any distribution of the form $G^{-1}(\mu)$ is called a *lifted distribution.*

Definition 3.12. For an affine space $A \subseteq \mathbb{F}_2^{nb}$ and a partial assignment $y \in \{0, 1, *\}^{nb}$, call y *extendable for* A if there exists $x \in A$ consistent with y

Definition 3.13. For an affine space $A \subseteq \mathbb{F}_2^{nb}$ and an extendable partial assignment $y \in \mathbb{F}_2^S$ (where $S \subseteq [nb]$) define $A_y \subseteq \mathbb{F}_2^{[nb] \setminus S}$ as follows:

$$A_y = \{\tilde{x} | (\tilde{x}, y) \in A\}$$

3.5 Linear algebraic facts about lifted spaces

In this subsection we import facts about closure and amortized closure proved by [EGI24], [AI25] and [EI25].

- **Safe set of vectors**

  Definition 3.14. (from [EGI24]) A set of vectors $V = \{v_1, v_2, \cdots, v_m\} \subseteq \mathbb{F}_2^{nb}$ is *safe* if for any k linearly independent vectors $w_1, w_2, \cdots, w_k \in \text{span}(S)$, $\text{supp}(w_1) \cup \text{supp}(w_2) \cup \cdots \cup \text{supp}(w_k)$ includes at least k distinct blocks.

- **Equivalent definition of *safe*:** Let

$$M = \begin{bmatrix} v_1 \\ \hline v_2 \\ \hline \cdots \\ \hline v_m \end{bmatrix} \in \mathbb{F}_2^{m \times nb}$$

Let $r = \text{rank}(M)$. V is nice iff there exist indices $c_1, c_2, \cdots, c_r \in [nb]$, each lying in different blocks, such that the set $\{Me_{c_1}, \cdots, Me_{c_r}\} \subseteq \mathbb{F}_2^m$ is linearly independent. (Me_j is the j-th column of M) The proof of equivalence of these two definitions can be found in Theorem 3.1 in [EGI24].

Fact 3.15. *Whether or not a set of vectors is safe depends only on their span. This is clear from the second equivalent definition.*

- **Safe affine spaces:**

Definition 3.16. Let $A \subseteq \mathbb{F}_2^{nb}$ be an affine space. Let $A = \{x|Mx = b\}$. Then, A is called a *safe affine space* if and only if the rows of M are safe. [1] By Fact 3.15, the choice of M does not affect this definition.

- **Deviolator:** For a subset of the blocks $S \subseteq [n]$ and a vector $v \in \mathbb{F}_2^{nb}$, define $v[\setminus S] \in \mathbb{F}_2^{(n-|S|)b}$ to be the projection of v on the coordinates of $[n] \setminus S$.

Definition 3.17. A subset $S \subseteq [n]$ is a deviolator for $V = \{v_1, v_2, \cdots, v_m\} \subseteq \mathbb{F}_2^{nb}$ if $\{v_1[\setminus S], v_2[\setminus S], \cdots, v_m[\setminus S]\} \subseteq \mathbb{F}_2^{(n-|S|)b}$ is a nice set.

- **Closure of a set of vectors:**

Definition 3.18. (from [EGI24]) Closure of a set $V = \{v_1, v_2, \cdots, v_m\}$ is the minimal deviolator for V. (It is known that this deviolator is unique, and also it depends only on $\text{span}(V)$ - Lemma 4.1 in [EGI24].)

- **Closure of an affine space:**

Definition 3.19. For an affine space A given by the set of equations $A = \{x|Mx = b\}$, define $\text{Cl}(A)$ to be the closure of the set of rows of M. [2]

- **Closure Assignment**

Definition 3.20. For an affine space A, a *closure assignment* y is any assignment to $\text{VARS}(\text{Cl}(A))$: $y \in \mathbb{F}_2^{\text{VARS}(\text{Cl}(A))}$.

- **Amortized Closure of a set of vectors**

[1] We emphasize that whenever we are talking about the safety of an affine space, we are talking about the safety of its set of defining equations.

[2] This does not depend on a specific choice of M.

A possible source of confusion could be that we defined closure for a set of vectors in item 4, and in item 5 we are defining the closure for an affine space in a different manner. In this paper, whenever we talk about the closure of a set of vectors V, we refer to the previous definition (item 4). Whenever we talk about the closure of an affine space A, we refer to this definition (item 5): the closure of its set of defining equations. This should not cause too much confusion: after the preliminaries section, we shall onlk talk about closures of affine spaces.

Definition 3.21. (from [AI25]) Let $V = \{v_1, v_2, \cdots, v_k\} \in \mathbb{F}_2^{nb}$. We define $\hat{\text{Cl}}(V) \subseteq [n]$ as follows: Let

$$M = \begin{bmatrix} v_1 \\ \hline v_2 \\ \hline \dots \\ \hline v_t \end{bmatrix}$$

Call a set of blocks $S = \{s_1, s_2, \cdots, s_k\} \subseteq [n]$ *acceptable* if there exist columns $c_1, c_2, \cdots, c_k$, such that c_j lies in block s_j and the set $\{Me_{c_1}, Me_{c_2}, \cdots, Me_{c_k}\}$ is linearly independent. The amortized closure of V, $\hat{\text{Cl}}(V)$, is the lexicographically largest acceptable set of blocks.

It is known that $\hat{\text{Cl}}(V)$ depends only on span(V) (Lemma 2.11 in [EI25])

- **Amortized Closure of An Affine Space**

  Definition 3.22. Let $A \subseteq \mathbb{F}_2^{nb}$ be an affine space; $A = \{x | Mx = b\}$. The amortized closure of A, $\hat{\text{Cl}}(A)$, is defined to be the amortized closure of the set of rows of M. This does not depend on a specific choice of M (Lemma 2.11 in [EI25])

Now we import some facts and lemmas about closure and amortized closure from [EGI24], [EI25] and [AI25].

Lemma 3.23. *If y is an extendable closure assignment of A, A_y is a safe affine subspace. (Follows from definition.)*

Lemma 3.24. *For any affine space, $Cl(A) \subseteq \hat{Cl}(A)$ (Lemma 2.15 in [EI25])*

Lemma 3.25. *If A, B are affine spaces with $B \subseteq A$, then $\hat{Cl}(A) \subseteq \hat{Cl}(B)$ (Corollary 2.19 in [EI25]) and $Cl(A) \subseteq Cl(B)$ (Lemma 4.2 in [EGI24])*

Lemma 3.26. *Let $V \subseteq \mathbb{F}_2^N$ be a set of vectors with $Cl(V) = S$. Let $W = V \cup \{e_{j,k} | j \in S, k \in [b]\}$. Then, $\hat{Cl}(V) = \hat{Cl}(W)$, $Cl(V) = Cl(W)$.*

Lemma 3.26 becomes clear once one examines the proof of Lemma 2.15 in [EI25] closely. For completeness we include a self-contained proof in Appendix A.

Lemma 3.27. *Let $V \subseteq W \subseteq \mathbb{F}_2^N$ be sets of vectors with $|W| = |V| + 1$. Then, $|\hat{Cl}(W)| \leq |\hat{Cl}(V)| + 1$, and moreover, if $|\hat{Cl}(W)| = |\hat{Cl}(V)| + 1$ then $Cl(W) = Cl(V)$ (Theorem 2.18 and Lemma 2.17 in [EI25]).*

We now state a useful corollary of the above.

Corollary 3.28. *Let $B \subseteq A$ be affine spaces such that $codim(B) = codim(A) + 1$ and $|\hat{Cl}(B)| = |\hat{Cl}(A)| + 1$. Let y be any extendable closure assignment for A. Then, A_y, B_y are both nice affine subspaces and $codim(B_y) = codim(A_y) + 1$.*

A proof of Corollary 3.28 is included in Appendix A.

4 Intuition and High-Level overview

At a high level, our proof combines the approaches of Alekseev and Itsykson [AI25], Efremenko and Itsykson [EI25] and Bhattacharya, Chattopadhyay and Dvorak [BCD24]. It does so by boosting the success probability of the ‘random walk with restart’ method of [AI25] by sampling inputs from a lifted distribution. The idea of using lifted distribution to do random walks appeared in [BCD24]. The bottleneck counting uses the notion of amortized closure instead of codimension of an affine space as done in [EI25]. However, combining these approaches requires significant new ideas – along with a new equidistribution lemma for gadgets with sufficiently small Fourier coefficients (Lemma 5.6). In this section we give a brief overview of how these approaches fit together.

Approach of Alekseev and Itsykson [AI25] We start by describing this approach which proved super-polynomial lower bounds for depth $n \log\log(n)$ Res($\oplus$). Their main idea is this: they take the CNF Ψ to be Tseitin contradiction over a $(n, \log(n), O(\log(n))$-expander lifted with an appropriate gadget; they assume we are given a size s Res-($\oplus$) refutation Π of Ψ, and they locate a path of length $n \log\log(n)$ in Π. They do this inductively: at Phase j, they locate a vertex v_j at depth $\Omega(nj)$. Given this vertex v_j, they show that as long as $\mathrm{codim}(A_{v_j})$ is not too large, there is another vertex v_{j+1} which is at distance $\Omega(n)$ from j. They show they can inductively find one more vertex as long as $j \le O(\log\log(n))$ - and this gives the depth lower bound.

Let us describe it in a bit more detail. Alekseev and Itsykson carefully choose a set of partial assignments in the unlifted world, $P \subseteq \{0,1,*\}^n$ with the idea that any partial assignment $\rho \in P$ leaves *some uncertainty* about which clause of the unlifted Tesitin formula would be falsified if one were to extend ρ at random to a full assignment.

In Phase j, Alekseev and Itsykson [AI25] have located a vertex v_j at depth $\Omega(jn)$. They want the codimension of A_{v_j}, the affine space that the proof Π associates with v_j, to be small ($\le O(j(\log(s/p))(b+1)^j)$, which is less than jn when j is small enough; p is a parameter we shall specify soon). Small co-dimension implies a small closure, i.e. $\mathrm{codim}(A_{v_j}) \ge |\mathrm{Cl}(A_{v_j})|$. We assume that variables in the unlifted world that correspond to blocks in $\mathrm{Cl}(A_{v_j})$ are revealed, but variables that correspond to blocks outside of the closure, i.e. in $[n] - \mathrm{Cl}(A_{v_j})$ are yet not revealed. Hence, Alekseev and Itsykson fix a closure assignment $y_j \in \mathbb{F}_2^{\mathrm{VARS}(\mathrm{Cl}(A_{v_j}))}$ such that $G(y_j)$ lies in P. Alekseev and Itsykson show (using a combinatorial argument) that the following holds when we uniformly sample a point $x \in A_{v_j} \cap C_y$ and follow the path of x from v_j for $\Theta(n)$ steps: with probability $\ge p$, after $\Theta(n)$ steps, the following holds: let w be the vertex reached. Let $\tilde{x} \in \mathbb{F}_2^{\mathrm{VARS}(\mathrm{Cl}(A_w))}$ be the restriction of x to the variables of $\mathrm{Cl}(A_w)$. Let $\rho \in \{0,1,*\}^n$ be the partial assignment that leaves all vars outside of $\mathrm{Cl}(A_w)$ free and $\rho|_{\mathrm{Cl}(A_w)} = G(\tilde{x})$. With probability at least p, ρ is in P, i.e. this ρ reveals little about where a potential falsified clause may be. For this combinatorial argument to work, it is essential that the current partial assignment, $G(y_j)$ lies in P and it does not fix too many bits: $|\mathrm{Cl}(A_{v_j})| \le O(n/\log(n))$.

One such w will be the next node, v_{j+1} - and the next closure assignment y_{j+1} could be anything in $\mathbb{F}_2^{\mathrm{VARS}(\mathrm{Cl}(A_w))}$ such that $G(y_{j+1}) \in P$ and y_{j+1} is extendible in A_w. The existence of such a y_{j+1} trivially follows as $x|_{\mathrm{VARS}(\mathrm{Cl}(A_w))}$ satisfies those requirements with non-zero probability. Note that all possible w's are at distance $\Omega(n)$ from v_j - so the only condition Alekseev and Itsykson need to maintain is that the codimension of A_w is not too high. They show the existence of such a w using a simple bottleneck argument: there exists a w such that a uniformly random node from $A_v \cap C_y$ reaches A_w with probability $\ge p/s$ as there are at most s many nodes at any given distance from node v_j. In particular, $|A_w| \ge \dfrac{p}{s}|A_v \cap C_y|$, which implies $\mathrm{codim}(A_w) \le (b+1)\mathrm{codim}(A_v) + \log(s/p) \le O((j+1)(\log(s/p))(b+1)^{j+1})$.

Let us briefly mention why this approach fails to go beyond depth $O(n \log\log n)$. Once $\mathrm{codim}(A_{v_j})$ exceeds $n/\log(n)$, the underlying combinatorial argument in [AI25] to get the next node fails. Hence, the depth lower bound obtained by this argument depends on the number of iterations till which $\mathrm{codim}(A_{v_j})$ is guaranteed to be less than $n/\log(n)$. In this case, there are two factors causing rapid growth of (the guaranteed upper bound on) $\mathrm{codim}(A_{v_j})$: first, at each step, the codimension of the next node can increase geometrically. Second, the success probability p in [AI25] is pretty low: around $2^{-O(n/\log(n))}$ - this also contributes to the growth of the valid upper bound on $\mathrm{codim}(A_{v_j})$

Improvement to depth $\Omega(N \log N)$: In 2025, Efremenko and Itsykson [EI25] bypassed the first barrier (of the codimension growing geometrically at each step) by introducing a new notion of progress other than the codimension: the amortized closure $\hat{\mathrm{Cl}}(A)$. Notice that the reason why the codimension was possibly growing geometrically in [AI25] is that fixing the bits of $\mathrm{Cl}(A_v)$ to y adds $b|\mathrm{Cl}(A_v)|$ more equations, which can be as large as $b \times \mathrm{codim}(A_v)$. One of the key observations in [EI25] is that if $|\hat{\mathrm{Cl}}(A_w)| = |\hat{\mathrm{Cl}}(v)| + k$, then $\Pr_{x\sim A_v \cap C_y}[x \in A_w] \leq 2^{-k}$. In other words, if $|\hat{\mathrm{Cl}}(A_w)| = |\hat{\mathrm{Cl}}(A_v)| + k$, among the equations defining A_w, there exist k linearly independent equations and moreover, these equations are also linearly independent from the equations of $A_v \cap C_y$ as the properties of amortized closure ensure $\hat{\mathrm{Cl}}(A_v) = \hat{\mathrm{Cl}}(A_v \cap C_y)$. Now, [EI25] runs the same argument again. This time, it yields the following recursion: $|\hat{\mathrm{Cl}}(A_{v_{j+1}})| \leq |\hat{\mathrm{Cl}}(A_{v_j})| + \log(\frac{s}{p})$, which prevents a geometric growth on the size of the amortized closure (as was happening with codimension earlier). This ensures that $|\hat{\mathrm{Cl}}(A_{v_j})| \leq O(j \log(s/p))$ at Phase j. This enables Efremenko and Itsykson [EI25] to find a vertex at depth $\Omega(N \log(N))$ assuming s was $\exp(N^{o(1)})$. However, the second barrier still remains: their success probability p is very small; around $2^{-n/\log(n)}$. Thus, this argument could not go beyond depth $N \log(N)$.

Our approach for depth $N^{2-\epsilon}$: One of the main contributions of this work is getting around this low success probability barrier. To do so, the starting point is the main idea of Bhattacharya, Chattopadhyay, and Dvorak [BCD24]. In [BCD24], the authors prove a separation between a restricted class of $\mathrm{Res}(\oplus)$ refutations (known as bottom-regular refutations) and general $\mathrm{Res}(\oplus)$ refutations. Their proof also employed a bottleneck argument, but instead of sampling from the uniform distribution, they were sampling from a lifted distribution. The key observation in [BCD24] was that if $g : \mathbb{F}_2^b \to \{0,1\}$ is an appropriate gadget, then for any lifted distribution $\bar{\mu}$ and any affine space A, $\Pr_{x\leftarrow\bar{\mu}}[x \in A] \leq 2^{-\Omega(\mathrm{codim}(A)/b)}$.

One might hope that the conditional version of such a statement is true: if $B \subseteq A$ are two affine spaces and $\tilde{\mu}$ is a lifted distribution, then $\Pr_{x\leftarrow\tilde{\mu}}[x \in B | x \in A] \leq 2^{-\Omega(\mathrm{codim}(B)-\mathrm{codim}(A))/b}$. If this were true, we could modify the proof of [AI25]: instead of sampling the input uniformly from $A_v \cap C_y$, we could sample from a lifted distribution tailored to our needs - which can hopefully boost the success probability. Unfortunately, such a statement cannot be true for any gadget, as the following counterexample shows.

Counterexample to conditional fooling

Let $t \in \mathbb{F}_2^b$ be a point, such that the first bit (wlog) is g-sensitive at t, i.e. $g(t) \neq g(t \oplus e_{\{1\}})$. WLOG, let $g(t) = 0$. The equations for A are as follows: for all $i \in [n], j \in [t] \setminus \{1\}, x_{ij} = t_j$. In B, we add the following extra equations: for all $i \in [n]$, $x_{i1} = t_1$. Let $\bar{\mu}$ be the uniform distribution on $G^{-1}(0^n)$. Then, even though $\mathrm{codim}(B) = \mathrm{codim}(A) + n$,

$$\Pr_{x\leftarrow\bar{\mu}}[x \in B | x \in A] = 1$$

Intuitively, the reason why conditional fooling does not happen in this counterexample is that A fixes too many linear forms in a block - and thus, when sampling from $G^{-1}(0^n) \cap A$, the distribution on each block is not controllable. One might imagine if the equations defining A do not concentrate too much on any single block, the distribution $G^{-1}(z) \cap A$ behaves more nicely. One notion of the equations defining A not concentrating on any single block is that A is a safe affine space. Indeed, it turns out that the conditional fooling conjecture is actually true when A and B are both safe affine spaces (Lemma 5.9) and the gadget has certain properties (gadget size $b = \Omega(\log(n))$ and all Fourier coefficients are exponentially small). And given this, it is not hard to show that lifted distributions fool amortized closure. Going back to the proof outline in [EI25], we shall show that if $\bar{\mu}$ is any lifted distribution and $|\hat{\mathrm{Cl}}(A_w)| = |\hat{\mathrm{Cl}}(A_v)| + k$, then $\Pr_{x \leftarrow \bar{\mu} \cap C_y}[x \in A_w | x \in A_v] \leq (3/4)^k$ (Lemma 5.1).

Lemma 5.1 provides us with the crucial flexibility for choosing any lifted distribution to boost the success probability of the random walk and re-start method. We exploit this to improve the sucess probability, from just $2^{-O(n/\log n)}$ in [EI25] to a constant independent of n in the following way: we continue with the base CNF being the Tseitin contradiction over a constant degree expander. We take the set of partial assignments to be (roughly) the same as the one taken by [AI25]. But now, given any partial assignment in P, we have to define a distribution on the unfixed variables whose lift will provide the required boost to success probability of the random walk. This part is new to our approach. The formal requirements are described in Section 7, called (p, q)-*PDT-hardness* – and the construction is described in Section 8. Analyzing the random walk on this new distribution requires further non-trivial technical ideas and we manage to prove that the success probability is indeed boosted from $2^{-n/\log(n)}$ to $1/3$ - this helps us remove the last $\log(n)$ factor, and we get a superpolynomial lower bound for depth-$N^{2-\epsilon}$ proofs in Res-$\oplus$ for any $\epsilon > 0$.

As a remark, we note that unlike [BCD24], [EI25], [AI25], we don't use any combinatorial properties of the gadget such as stifling. The property of the gadget we use is exponentially small $||\hat{g}||_\infty$. The proof of the conditional fooling lemma is significantly more involved than the proof of the vanilla fooling lemma in [BCD24]; we need to employ an exponential sum argument, whereas the vanilla fooling lemma has a simple combinatorial proof using stifling.

It also seems that N^2 is the best depth lower bound one can hope for using the random walk with restarts technique: at each phase, the path taken can have length at most n, and the amortized closure will increase by at least 1. We obtain a depth lower bound that is arbitrarily close to N^2. One would require new ideas to obtain depth $\omega(N^2)$-lower bounds.

5 Conditional fooling lemma

Throughout this section, assume the gadget $g : \mathbb{F}_2^b \to \mathbb{F}_2$ has the following properties.

- $||\hat{g}||_\infty \leq 2^{-\alpha b}$ for some constant $\alpha > 0$. In other words, for all $S \subseteq [b]$, $|\hat{g}(S)| = \left|\mathrm{E}[g(x)(-1)^{\sum_{i \in S} x_i}]\right| \leq 2^{-\alpha b}$
- The gadget size is $b(n) = \dfrac{250}{\alpha} \log(n)$.

In this section, we will establish a key result that shows that lifted distributions *fool* amortized closure (Lemma 5.1).

In the following, we state a fact that will be, in some sense, a significant generalization of the following simple, well known fact: If $B \subseteq A \subseteq \mathbb{F}_2^{nb}$ are two affine spaces, then $Pr_{x \sim A}[x \in B] \leq 2^{\mathrm{codim}(A) - \mathrm{codim}(B)}$. This fact was generalized recently by Efremenko and Itsykson [EI25]. Let y be an extendible assignment to the variables in closure of A, i.e. $\mathrm{Cl}(A)$. Then, Lemma 5.1 of [EI25], that they point out is their key lemma for improving the lower bound on resolution over parities, shows the following:

$$\Pr_{x \sim A \cap C_y}[x \in B] \leq 2^{|\hat{\mathrm{Cl}}(A)| - |\hat{\mathrm{Cl}}(B)|}. \tag{5.1}$$

Note here we cannot hope to work with co-dimension of B and A since, for example, B might precisely be those elements in A that extend y which is an affine space and $\mathrm{codim}(B) - \mathrm{codim}(A) = |y|$. The argument in [EI25] uses a convenient property of amortized closure, combined with simple linear algebra. In another direction, Bhattacharya, Chattopadhyay and Dvorak [BCD24] showed the following: if g is a gadget with safe properties, then the following is true for every $z \in \{0,1\}^n$:

$$\Pr_{x \sim G^{-1}(z)}[x \in B] \leq 2^{-\Omega(\mathrm{codim}(B)/b)} \tag{5.2}$$

Below, we prove our main lemma which has the features of both (5.1) and (5.2).

Lemma 5.1. *Let $B \subseteq A \subseteq \mathbb{F}_2^{nb}$ be affine subspaces such that $|\hat{Cl}(B)| = |\hat{Cl}(A)| + k$. Let $y \in \mathbb{F}_2^{VARS(Cl(A))}$ be an extendable closure assignment of A, and let $z \in \mathbb{F}_2^n$ be a point such that $G(y) = z|_{Cl(A)}$. Then,*

$$Pr_{x \sim G^{-1}(z) \cap C_y}[x \in B | x \in A] \leq \left(\frac{3}{4}\right)^k$$

Currently, we do not know of a short argument to prove this. We prove it here in steps, establishing some equidistribution properties of gadgets with small Fourier coefficients wrt *safe* affine spaces that seem independently interesting.

Lemma 5.2. *Let $A \subseteq \mathbb{F}_2^{nb}$ be a safe affine subspace with $codim(A) = m$. Let $z \in \mathbb{F}_2^n$ be any target assignment. Then,*

$$Pr_{x \sim \mathbb{F}_2^{nb}}[x \in A \wedge G(x) = z] \in \left[\frac{1 - o(n^{-100})}{2^{m+n}}, \frac{1 + o(n^{-100})}{2^{m+n}}\right].$$

Proof. Let M be a matrix for the equations defining A. Since A is safe, there exist m blocks such that one can choose one column from each block, such that those columns are linearly independent. WLOG (for notational convenience) assume those blocks are $1, 2, \cdots, m$, and from block j we choose column a_j.

We first rewrite the system of equations in a more convenient form. Since the matrix M restricted to column set $S = \{(j, a_j) | 1 \leq j \leq m\}$ is invertible, we can perform row operations on M so that the submatrix $M_{[m],S}$ becomes I_m. Let ℓ_i denote the i-th row of this modified matrix. Thus, for every $i \in [m]$, there exists a $c \in [b]$ such that l_i has a non-zero entry at coordinate (i, c), and for every $i' \neq i$, $l_{i'}$ has a zero entry at coordinate (i, c). An easy but crucial consequence of this is the following.

Observation 5.3. For every subset $T \subseteq [m]$, the vector $\sum_{j \in T} \ell_j$ has a non-zero coordinate in the j-th block for each $j \in T$.

Suppose the system of equations in this basis is

$$\begin{aligned} \langle \ell_1, x \rangle &= c_1 \\ \langle \ell_2, x \rangle &= c_2 \\ &\cdots\cdots \\ \langle \ell_m, x \rangle &= c_m \end{aligned}$$

Notation: for an assignment $x \in \mathbb{F}_2^{nb}$, we denote by $x(i) \in \mathbb{F}_2^b$ the restriction of x to the i'th block.

Let $p := \Pr_{x \sim \mathbb{F}_2^{nb}}[x \in A \wedge G(x) = z]$. We have

$$p = \mathrm{E}_x \left(\prod_{j=1}^{n} \left(\frac{1 + (-1)^{g(x(i)) + z_i}}{2} \right) \prod_{j=1}^{m} \left(\frac{1 + (-1)^{\ell_j(x) + c_j}}{2} \right) \right)$$

Expanding the RHS, we get the following expression:

$$p - \frac{1}{2^{n+m}} = \sum_{\substack{S \subseteq [n] \\ T \subseteq [m] \\ S \cup T \neq \phi}} \mathrm{E}_x \left[\frac{(-1)^{\sum\limits_{i \in S} (g(x(i)) + z_i) + \sum\limits_{j \in T} (\ell_j(x) + c_j)}}{2^{n+m}} \right]$$

For $S \subseteq [n], T \subseteq [m]$ let $f_{S,T}(x) \coloneqq (-1)^{\sum\limits_{i \in S} g(x(i)) + \sum\limits_{j \in T} \ell_j(x)}$ and $u_{S,T} \coloneqq \sum\limits_{i \in S} z_i + \sum\limits_{j \in T} c_j$. We have

$$p - \frac{1}{2^{n+m}} = \frac{1}{2^{n+m}} \sum_{\substack{S \subseteq [n] \\ T \subseteq [m] \\ S \cup T \neq \phi}} (-1)^{u_{S,T}} \mathrm{E}_x[f_{S,T}(x)]$$

We start by showing that $\mathrm{E}_x[f_{S,T}(x)]$ vanishes unless $T \subseteq S$. This is where we use the safety of A.

Claim 5.4. If $T \not\subseteq S$, $\mathrm{E}_x[f_{S,T}(x)] = 0$

Proof. Let $u \in T \setminus S$. By Observation 5.3, there exists a coordinate k in the u-th block on which $\sum\limits_{j \in T} \ell_j$ is non-zero: $\sum\limits_{j \in T} (\ell_j)_{(u,k)} = 1$. Since $u \notin S$, this coordinate does not affect $\sum\limits_{i \in S} g(x(i))$. So we have that for all x, $f_{S,T}(x) = -f_{S,T}(x \oplus e_{u,k})$. Therefore, exactly half of the x's have $f_{S,T}(x) = 1$ and the result follows. □

It now suffices to bound the terms where $T \subseteq S$. We do this using the fact that all Fourier coefficients of g are small.

Claim 5.5. If $T \subseteq S$, $|\mathrm{E}_x[f_{S,T}(x)]| \leq \frac{1}{2^{\alpha b |S|}}$

Proof. Let $g^{\oplus S} : \mathbb{F}_2^{b \times |S|} \to \mathbb{F}_2$ be the XOR of $|S|$ disjoint copies of g; for $y \in \mathbb{F}_2^{b \times |S|}$,

$$g^{\oplus S}(y) = \left(\sum_{i \in S} g(y(i)) \right) \pmod 2$$

Note that $||\widehat{g^{\oplus S}}||_\infty = (||\hat{g}||_\infty)^{|S|} \leq 2^{-\alpha b |S|}$. Therefore,

$$|\mathrm{E}_x[f_{S,T}(x)]| = \left| \widehat{g^{\oplus S}} \left(\mathrm{supp} \left(\sum_{j \in T} l_j \right) \right) \right| \leq 2^{-\alpha b |S|}.$$

□

Now, we upper bound the magnitude of the error as follows. If $|S| = k$, Claim 5.4 implies that there are at most 2^k possible values of T for which $\mathrm{E}_x[f_{S,T}(x)] \neq 0$. Claim 5.5 implies that the magnitude of each of these terms is at most $\frac{1}{2^{\alpha bk}}$. Thus, in our setting of $b = \frac{250}{\alpha}\log(n)$, we get that

$$\left|p - \frac{1}{2^{n+m}}\right| = \frac{1}{2^{n+m}}\left|\sum_{\substack{S\subseteq[n]\\T\subseteq[m]\\S\cup T\neq\phi]}}(-1)^{u_{S,T}}\mathrm{E}_x[f_{S,T}(x)]\right|$$

$$\leq \frac{1}{2^{n+m}}\sum_{k=1}^{n}\binom{n}{k}2^k\frac{1}{2^{\alpha bk}}$$
$$\leq \frac{1}{2^{n+m}}o(n^{-100})$$

This completes the proof. □

Corollary 5.6. *If A is a safe affine space, for all $z \in \mathbb{F}_2^n$,*

$$Pr_{x\sim A}[G(x) = z] \in \left[1 \pm o(n^{-100})\right]\frac{1}{2^n}$$

Proof.
$$\mathrm{Pr}_{x\sim A}[G(x) = z] = \frac{\mathrm{Pr}_x[x \in A \wedge G(x) = z]}{\mathrm{Pr}_x[x \in A]}$$

The denominator is $2^{-\mathrm{codim}(A)}$, and to estimate the numerator use Lemma 5.2. □

We will now show that the set of pre-images of an arbitrary $z \in \{0,1\}^n$, are approximately equidistributed among the various translates of a safe affine space in the lifted world.

Lemma 5.7. *Let $A \subseteq \mathbb{F}_2^{nb}$ be a safe affine space with codimension m, and let $z \in \mathbb{F}_2^n$ be a target point. Then,*

$$Pr_{x\sim G^{-1}(z)}[x \in A] \in \left[\frac{1-o(n^{-90})}{2^m}, \frac{1+o(n^{-90})}{2^m}\right]$$

Proof. Let $A_1 = A, A_2, \cdots, A_M$ be the $M = 2^m$ translates of A. Let $S_j = G^{-1}(z) \cap A_j$. Lemma 5.6 implies $\frac{|S_j|}{|A|} \in \left[\frac{1-o(n^{-100})}{2^n}, \frac{1+o(n^{-100})}{2^n}\right]$ for all j. We have

$$\mathrm{Pr}_{x\sim G^{-1}(z)}[x \in A] = \frac{|S_1|}{\sum_j |S_j|} \in \left[\frac{1-o(n^{-100})}{1+o(n^{-100})}\times\frac{1}{2^m}, \frac{1+o(n^{-100})}{1-o(n^{-100})}\times\frac{1}{2^m}\right]$$
$$= \left[\frac{1-o(n^{-90})}{2^m}, \frac{1+o(n^{-90})}{2^m}\right]$$

□

Using the above, we show below that if $B \subset A$ are two safe affine spaces, then B cannot significantly distinguish the distributions $x \sim (G^{-1}(z) \cap A)$ and $x \sim A$.

Lemma 5.8. *Let $B \subseteq A \in \mathbb{F}_2^{nb}$ be safe affine subspaces such that $codim(B) = codim(A) + 1$. Let $z \in \mathbb{F}_2^n$ be any point. Then,*

$$Pr_{x \sim G^{-1}(z)}[x \in B | x \in A] \leq \frac{1}{2} + o(n^{-50})$$

Proof. Let $m = \text{codim}(A)$. Lemma 5.7 implies $\Pr_{x \sim G^{-1}(z)}[x \in A] \geq \frac{1 - o(n^{-90})}{2^m}$ and $\Pr_{x \sim G^{-1}(z)}[x \in B] \leq \frac{1 + o(n^{-90})}{2^{m+1}}$ Thus,

$$\Pr_{x \sim G^{-1}(z)}[x \in B | x \in A] = \frac{\Pr_{x \sim G^{-1}(z)}[x \in B]}{\Pr_{x \sim G^{-1}(z)}[x \in A]} \leq \frac{1 + o(n^{-90})}{1 - o(n^{-90})} \times \frac{1}{2} \leq \frac{1}{2} + o(n^{-50})$$

□

The structure of safe affine spaces that we have discovered so far allows us to say the following about any two arbitrary affine spaces that are not necessarily safe.

Lemma 5.9. *Let $B \subseteq A \in \mathbb{F}_2^{nb}$ be affine spaces such that $|\hat{Cl}(B)| = |\hat{Cl}(A)| + 1$ and $codim(B) = codim(A) + 1$. Let y be an extendable closure assignment for A, and let $z \in \mathbb{F}_2^n$ be a point such that $z|_{Cl(A)} = G(y)$. Then,*

$$Pr_{x \sim G^{-1}(z) \cap C_y}[x \in B | x \in A] \leq \frac{1}{2} + o(n^{-50})$$

Proof. Let $z = (G(y), w)$. Rewrite the desired probability expression as

$$\Pr_{\tilde{x} \sim G^{-1}(w)}[x \in B_y | x \in A_y]$$

By Corollary 3.28, A_y, B_y are both safe affine subspaces, and $\text{codim}(B_y) = \text{codim}(A_y) + 1$. Now the result follows from Lemma 5.8. □

An easy corollary is that the result still holds if we condition only on a subset of the blocks in $\text{Cl}(A)$ instead of all the blocks in $\text{Cl}(A)$.

Corollary 5.10. *Let $B \subseteq A \in \mathbb{F}_2^{nb}$ be affine spaces such that $|\hat{Cl}(B)| = |\hat{Cl}(A)| + 1$ and $codim(B) = codim(A) + 1$. Let $S \subseteq Cl(A)$ and let $y \in \mathbb{F}_2^{VARS(S)}$ be a partial assignment. Let $z \in \mathbb{F}_2^n$ be a point such that $z|_S = G(y)$ and $G^{-1}(z) \cap C_y \cap A \neq \phi$. Then,*

$$Pr_{x \sim G^{-1}(z) \cap C_y}[x \in B | x \in A] \leq \frac{1}{2} + o(n^{-50})$$

Proof. Sampling x from $G^{-1}(z) \cap C_y$ can be done as follows: first sample $y^{(1)} \in \mathbb{F}_2^{\text{VARS}(\text{Cl}(A))} \cap C_y$ according to $G^{-1}(z)$, then sample x from $G^{-1}(z) \cap C_{y^{(1)}}$. For each possible $y^{(1)}$ use Lemma 5.9 to upper bound the conditional probability of lying in B conditioned on $y^{(1)}$. Formally, let $\mathcal{D}$ denote the distribution of $x|_{\text{Cl}(A)}$ as $x \sim G^{-1}(z) \cap C_y$. Then,

$$\begin{aligned}\Pr_{x \sim G^{-1}(z) \cap C_y}[x \in B | x \in A] &= \text{E}_{y^{(1)} \leftarrow \mathcal{D}}[\Pr_{x \sim G^{-1}(z) \cap C_{y^{(1)}}}[x \in B | x \in A]] \\ &\leq \frac{1}{2} + o(n^{-50})\end{aligned}$$

□

Now we prove the final result of this section.

Proof. of Lemma 5.1 Let $B = W_0 \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_{l-1} \subseteq W_l = A$ be a sequence of affine subspaces such that $\text{codim}(W_j) = \text{codim}(W_{j+1}) + 1$. We have

$$\Pr_{x \sim G^{-1}(z) \cap C_y}[x \in B | x \in A] = \prod_{j=0}^{l-1} \Pr_{x \sim G^{-1}(z) \cap C_y}[x \in W_j | x \in W_{j+1}]$$

We assume there exists a point in $G^{-1}(z) \cap C_y \cap B$ (as otherwise the conditional probability is 0), so in particular, for all j there exists a point in $G^{-1}(z) \cap C_y \cap C_j$.

By Lemma 3.27 there exist k indices $j \in \{0, 1, \cdots, l-1\}$ such that $|\hat{\text{Cl}}(W_j)| = |\hat{\text{Cl}}(W_{j+1})| + 1$. Note that $\text{Cl}(A) \subseteq \text{Cl}(W_{j+1})$ by Lemma 3.25. Invoking Corollary 5.10 for each such index j, where W_j plays the role of B and W_{j+1} that of A, we have

$$\Pr_{x \sim G^{-1}(z) \cap C_y}[x \in W_j | x \in W_{j+1}] \leq \frac{3}{4}$$

So, in the product $\prod_{j=0}^{l-1} \Pr_{x \sim G^{-1}(z) \cap C_y}[x \in W_j | x \in W_{j+1}]$, at least k terms are $\leq 3/4$. The result follows. □

6 Description of CNF

The CNF we shall use is the Tseitin contradiction over an expander graph, lifted with an appropriate gadget. Let $G = (V, E)$ be a $(|V|, d, \lambda < 1/1000)$ expander with $|V|$ odd and $d = O(1)$. The base CNF Φ has variables $z_{u,v}$ for $(u, v) \in E$. For each $v \in V$ we express the constraint $\sum_{(v,w) \in E} z_{v,w} \equiv 1 \pmod 2$ using $2^d = O(1)$ clauses. This system is unsatisfiable because adding up all the equations yields $0 \equiv 1 \pmod 2$.

The property of G we shall use is isoperimetric expansion (which follows from Cheeger's inequality [Che71]):

Lemma 6.1. *For any $S \subseteq V$, the cut $E(S, V \setminus S)$ has at least $\frac{d}{5} \min(|S|, n - |S|)$ edges.*

Explicit constructions of such graphs were provided in [LPS88] and [Mar73].

We lift Φ with an appropriate gadget. We will take the gadget $g : \mathbb{F}_2^b \to \mathbb{F}_2$ to have the properties mentioned in section 5. For convenience of the reader we restate the properties here.

- $||\hat{g}||_\infty \leq 2^{-\alpha b}$ for some constant $\alpha > 0$. In other words, for all $S \subseteq [b]$, $|\hat{g}(S)| = \left|\text{E}[g(x)(-1)^{\sum_{i \in S} x_i}]\right| \leq 2^{-\alpha b}$
- The gadget size is $b(n) = \frac{250}{\alpha} \log(n)$.

A concrete instantiation of g is the Inner Product function $g = \text{IP}$:

$$\text{IP}(x_1, x_2, \cdots, x_{b/2}, y_1, y_2, \cdots, y_{b/2}) = (x_1 y_1 + \cdots + x_{b/2} y_{b/2}) \pmod 2$$

Theorem 6.2. *For $g = IP$, $||\hat{g}||_\infty \leq 2^{-b/2}$*

In the case of $g = \text{IP}$ we get $\alpha = 1/2$ and $b(n) = 500 \log(n)$. Throughout the rest of the paper, we shall assume the gadget g has all Fourier coefficients $\leq 2^{-\alpha b}$ in magnitude and $b(n) = \frac{250}{\alpha} \log(n)$.

The CNF for which we will prove our depth restricted lower bounds is $\Psi = \Phi \circ g$. Note that the number of clauses in Ψ is $N^{O(1)}$.

7 The Utility of (p, q)-PDT Hardness

Alekseev and Itsykson [AI25] introduced the 'random walk with restarts' approach to prove superlinear lower bounds on depth of Res($\oplus$) proofs of small size. To analyze their random walk with restarts, [AI25] uses certain elaborate games. We find it more convenient to analyze random walks using the language of decision trees. In particular, this allows us naturally to bring in the notion of a hard distribution that seems crucial to boost the success probability of our random walk with restart significantly, all the way from $2^{-n/\log(n)}$ to a constant. In this section, we formalize our notion which we call (p, q)-PDT hardness. We point out that our notion here is a significant refinement of the ideas of Bhattacharya, Chattopadhyay and Dvorák [BCD24] where as well random walks on lifted distributions were analyzed, but without restarts.

We first set up some notation to define our hardness notion. For a parity decision tree T and a point x, define the affine subspace $A_x(T)$ to be the one corresponding to the set of inputs y that traverse the same path in T as x does. More formally, $A_x(T)$ is defined as follows: suppose on input x, T queries the linear forms $\ell_1, \cdots, \ell_d$ and gets responses $c_1, c_2, \cdots, c_d$ respectively. Then, $A_x(T) = \{y | \langle \ell_j, y\rangle = c_j \forall j \in [d]\}$.

We are ready now to introduce the notion of a hard set of partial assignments that will abstract our requirements for finding a deep node in the proof DAG.

Definition 7.1. Let Φ be a CNF formula on n variables. A non-empty set of partial assignments $P \subseteq \{0, 1, *\}^n$ is (p, q)-*PDT-hard* for Φ if the following properties hold:

- **No falsification:** No partial assignment in P falsifies any clause of Φ.
- **Downward closure:** If $\rho \in P$ and $\tilde{\rho}$ is obtained from ρ by unfixing some of the bits set in ρ, then $\tilde{\rho} \in P$
- **Hardness against parity decision trees:** Let $A \subseteq \mathbb{F}_2^{nb}$ be an affine space with $|\hat{\mathrm{Cl}}(A)| \le p$. Let $y \in \mathbb{F}_2^{\mathrm{VARS}(\mathrm{Cl}(A))}$ be an extendable closure assignment for A such that $\alpha = G(y) \in P$. Then, there exists a distribution $\mu = \mu(\alpha)$ on $\mathbb{F}_2^n$ such that the following properties hold:
  1. $z|_{\mathrm{Cl}(A)} = \alpha$ for all $z \in \mathrm{supp}(\mu)$
  2. Let T be any parity decision tree (with input nb bits) of depth $\le q$. For any x, define $\tilde{A}(x) = A_T(x) \cap A \cap C_y$.With probability $\ge 1/3$, as x is sampled from $G^{-1}(\mu) \cap A \cap C_y$, it holds that $G(x|_{\mathrm{Cl}(\tilde{A}(x))}) \in P$ [3].

The CNF Φ is (p, q)-*PDT-hard* if it admits a non-empty (p, q)-PDT-hard set of partial assignments.

We now state the main result of this section that shows (p, q)-PDT-hardness of a CNF is sufficient to get us good lower bound on depth of a refutation of the lifted formula, assuming the size of the refutation is small.

Theorem 7.2. *Let Φ be a CNF on n variables having a non-empty (p, q)-PDT-hard set of partial distributions. Then, any Res($\oplus$) refutation of $\Phi \circ g$ of size s must have depth at least* $\Omega\left(\frac{pq}{\log(s)}\right)$.

[3] $G^{-1}(\mu) \cap A \cap C_y$ is non-empty by Lemma 5.6 applied on the nice affine space A_y

To prove the above, we will first establish the following lemma. This lemma essentially tells us that as long as we are at a node whose associated affine space satisfies some convenient properties, we are assured to find another node at a distance q from our starting node whose corresponding affine space continues to have reasonably convenient properties.

Lemma 7.3. *Suppose Φ has a non-empty (p,q)-PDT-hard set of partial assignments P. Let Π be a Res($\oplus$) refutation of $\Phi \circ g$ of size s. Let v be a node in Π such that $|\hat{Cl}(A_v)| \leq p$, and let $y \in \mathbb{F}_2^{VARS(Cl(A_v))}$ be an extendable closure assignment for A_v such that $G(y) \in P$. Then, there exists another node w in Π such that:*

1. *There exists a length q path from v to w in Π.*
2. *There exists an extendable closure assignment for A_w, $\tilde{y}$, such that $G(\tilde{y}) \in P$.*
3. *$|\hat{Cl}(A_w)| \leq |\hat{Cl}(A_v)| + 2\log(s)$*

Proof. Let $\mu = \mu(\alpha)$ be the hard distribution guaranteed to exist by the definition of (p,q)-PDT-hardness, where $\alpha = G(y)$. Let T be the following parity decision tree: on any input x, it simulates the queries made by Π starting from node v for q steps. For any $x \in A_v$, define $\text{END}_q(x)$ to be the node of Π reached by x starting from v after q steps. (In case Π on x reaches a leaf within q steps starting from v, define $\text{END}_q(x)$ to be that leaf.)

We have $A_T(x) \cap A_v \subseteq A_{\text{END}_q(x)}$. Let $\text{GOOD} = \{x | G(x|_{\text{Cl}(\tilde{A}_v(x))}) \in P\}$ (recall, $\tilde{A}_v(x) = A_T(x) \cap A_v \cap C_y$). The definition of (p,q)-PDT-hardness guarantees that $\Pr_{x \leftarrow G^{-1}(\mu) \cap A \cap C_y}[x \in \text{GOOD}] \geq 1/3$.
Let $\mathcal{N} = \{\text{END}_q(x) | x \in \text{GOOD}\}$. Note that since no assignment in P falsifies any clause of Φ, no vertex in $\mathcal{N}$ is a leaf - and therefore, there is a length q walk from v to w for all $w \in \mathcal{N}$ (i.e., the parity decision tree does not terminate before q queries if $x \in \text{GOOD}$). Also, $A_T(x) \cap A_v \cap C_y \subseteq A_{\text{END}_q(x)}$, so $\text{Cl}(A_{\text{END}_q(x)}) \subseteq \text{Cl}(A_T(x) \cap A_v \cap C_y)$, so $x \in \text{GOOD}$ implies $G(x|_{\text{Cl}(\text{END}_q(x))}) \in P$ (since P is downward closed). Thus, properties (i) and (ii) are satisfied for all $w \in \mathcal{N}$. To complete the proof, we have to find a $w \in \mathcal{N}$ such that $|\hat{\text{Cl}}(A_w)| \leq |\hat{\text{Cl}}(A_v)| + 2\log(s)$.
Since $|\mathcal{N}| \leq s$, there exists a $w \in \mathcal{N}$ such that $\Pr_{x \leftarrow G^{-1}(\mu) \cap C_y \cap A}[\text{END}_q(x) = w] \geq \dfrac{1}{3s}$. In particular, this implies

$$\Pr_{x \leftarrow G^{-1}(\mu) \cap C_y}[x \in A_w | x \in A] \geq \frac{1}{3s}$$

Lemma 5.1 then implies $|\hat{\text{Cl}}(A_w)| \leq |\hat{\text{Cl}}(A_v)| + 2\log(s)$

□

Now we are ready to prove our main result for this section, by repeatedly making use of Lemma 7.3.

Proof of Theorem 7.2. Let Π be a Res($\oplus$) refutation of Φ. We shall inductively find vertices $v_1, v_2, \ldots, v_j$ in Π for $j \leq \dfrac{p}{2\log(s)}$ such that:

- $\text{depth}(v_j) \geq jq$
- $|\hat{\text{Cl}}(A_{v_j})| \leq 2j\log(s)$
- There exists an extendable closure assignment y_j for A_{v_j} such that $G(y_j) \in P$

For $j = 0$ we pick the root. To get v_{j+1} we apply Lemma 7.3 to v_j. We can continue this way as long as $|\hat{\text{Cl}}(A_{v_j})| \leq p$. Hence, we do this for $j = \left\lfloor \frac{p}{2\log s} \right\rfloor$ many steps. In the end, we get a node at depth $\Omega\left(\dfrac{pq}{\log(s)}\right)$. □

8 Proving (p, q)-PDT hardness

Our goal in this section is to show that the Tseitin contradiction over an expander graph meets the requirements of $(\Omega(n), \Omega(n))$-PDT hardness as specified in Definition 7.1 in the previous section.

Note that (p, q)-PDT hardness is a notion that measures the hardness of a set of partial assignments against parity decision trees. We start by proving that vanilla Tseitin tautologies defined over constant-degree expanders graphs (as opposed to being lifted by a gadget), already satisfy a weaker notion of hardness that we call (p, q)-DT hardness that is effective against ordinary decision trees operating in the unlifted world. This will allow us to introduce many of the ideas more cleanly that will then be re-used in the more involved lifted world of parity decision trees. After establishing (p, q)-DT hardness in the unlifted setting, we shall sketch how to modify the argument to prove the original requirement of $(\Omega(n), \Omega(n))$-PDT hardness in the lifted setting. For completeness, a self-contained and direct proof of the original requirement of $(\Omega(n), \Omega(n))$-PDT hardness is presented in Appendix B.

8.1 Hardness Against Ordinary Decision Trees

We define the analogue of (p, q)-PDT-hardness in the unlifted setting.

Definition 8.1 ((p, q)-DT hardness)**.** For a CNF Φ on n variables, call a set of partial assignments $P \subseteq \{0, 1, *\}^n$ to be (p, q)-DT-hard if the following hold:

- **No falsification:** No partial assignment $\rho \in P$ falsifies any clause of Φ.
- **Downward closure:** For any $\rho \in P$ and any $j \in [n]$, if $\tilde{\rho}$ is obtained by setting $\rho(j) \leftarrow^*$, then $\tilde{\rho} \in P$
- **Hard for decision trees:** For any $\rho \in P$ which fixes at most p variables, there exists a distribution μ on the assignment to unfixed variables such that the following holds:
  - Let T be a decision tree of depth q querying the unfixed variables. If we sample an assignment to the unfixed variables from μ and run T for q steps, the partial assignment we see in the end also lies in P with probability $\geq 1/3$.

The CNF Φ is (p, q)-DT hard if it admits a non-empty set of (p, q)-DT-hard partial assignments.

Recall that we are working with the base CNF to be a Tseitin contradiction over an expander graph, as defined in Section 6. The main theorem of this subsection is that Tseitin contradiction over an expander is $(\Omega(n), \Omega(n))$-DT hard.

Theorem 8.2. *Let Φ be the Tseitin contradiction over a $(|V|, d, \lambda < 1/1000)$ expander (with $|V|$ odd). Then, Φ is $(n/2000, n/2000)$-DT-hard – i.e., there exists a non-empty $(n/2000, n/2000)$-DT-hard set of partial assignments for Φ.*

8.1.1 Choosing the set of partial assignments

We define a partial assignment to the edges of our graph to be *valid* below. The set $P \subseteq \{0, 1, *\}^n$ will be the set of valid partial assignments.

Definition 8.3. Let $\rho \in \{0, 1, *\}^E$ be a partial assignment. We define the criteria for checking if ρ is *valid* as follows: let $S = \{e | e \text{ has been fixed by } \rho\}$. For each $v \in V$, define $f_\rho(v) = 1 + \sum_{(v,w) \in S} \rho(v, w)$ (i.e., $f_\rho(v)$ denotes the parity of the unfixed edges incident to v in order to satisfy the original degree constraint for v). Then, ρ is said to be *valid* if the following conditions are satisfied:

1. There exists exactly one connected component C in $(V, E \setminus S)$ such that $\sum_{v \in C} f_\rho(v) = 1$.
   We call this component the *odd* component and every other component is called *even*.
2. The size of the odd connected component, $|C|$, is more than $n/2$.

We now show that the set P of all valid partial assignments is indeed (p,q)-DT-hard. We begin by showing below that the first two properties for being (p,q)-DT-hard are satisfied.

Lemma 8.4. *The set of partial assignments P satisfies the conditions **Downward Closure** and **No falsification** for Φ (as defined in Definition 8.1).*

Proof. Both properties are straightforward to verify.

- **No falsification:** In order to falsify any clause, ρ has to fix all edges of some vertex. In that case, that vertex is an isolated connected component in $(V, E \setminus S)$ and the total f_ρ in that component is 1. However, the first condition stipulates that there is exactly one connected component whose total f_ρ is odd, and that component has size more than $n/2$.
- **Downward closure:** Let $\rho \in P$, and let $\tilde{\rho}$ be obtained from ρ by setting $\rho(e) =^*$ for some $e \in E$ that was fixed by ρ. There are three cases.
  1. $e = (a,b)$ bridges the largest component C with some other component W. Wlog, $a \in C$ and $b \in W$. Let $S' = S \setminus \{e\}$. Let the new expanded connected component be $C' = C \cup W$. Note that W forms a connected component in $(V, E \setminus S)$ and therefore $\sum_{v \in W} f_\rho(v) = 0 \pmod 2$. For each $v \neq a, b$, $f_{\tilde{\rho}}(v) = f_\rho(v)$, and $f_{\tilde{\rho}}(a) = \rho(e) + f_\rho(a) \pmod 2$, $f_{\tilde{\rho}}(b) = \rho(e) + f_\rho(b) \pmod 2$. We have $|C'| \geq |C| \geq n/2$, so all we need to verify is that $\sum_{e \in C'} f_{\tilde{\rho}}(v) \equiv 1 \pmod 2$.
  $$\sum_{v \in C \cup W} f_{\tilde{\rho}}(v) = \sum_{v \in C} f_\rho(v) + \rho(e) + \sum_{v \in W} f_\rho(v) + \rho(e) = 1 + 0 = 1 \pmod 2$$
  2. $e = (a,b)$ bridges two components U and W, none of which is C. Very similar argument as above shows that
  $$\sum_{v \in U \cup W} f_{\tilde{\rho}}(v) = \sum_{v \in U} f_\rho(v) + \rho(e) + \sum_{v \in W} f_\rho(v) + \rho(e) = 0 + 0 = 0 \pmod 2$$
  Thus, C remains the unique odd connected component.
  3. $e = (a,b)$ does not bridge two different components. It is simple to verify in this case that the parity of all components remain unchanged.

□

Now we come to the final property: hardness for decision trees. First, we prove an easy but crucial lemma:

Lemma 8.5. *Let $G = (V,E)$ be a connected undirected graph and let $g \in \mathbb{F}_2^V$ a vector. Let $T \subseteq E$ be a spanning tree, and let $h \in \mathbb{F}_2^{E \setminus T}$ be an assignment to the edges not in T. Let $v \in V$ be a vertex. There exists a unique assignment $\tilde{h} \in \mathbb{F}_2^E$ such that $\tilde{h}$ extends h and $\sum_{w \in N(u)} \tilde{h}(u,w) = g(u)$ for all $u \neq v$.*

Proof. We construct $\tilde{h}$ as follows: convert T to a rooted tree by making v the root and then process the vertices bottom up, starting at the leaves of T. When vertex u is being processed, all edges in the subtree of u have been assigned. Then, exactly one edge incident to u is kept unfixed (the edge $(u, \text{parent}[u])$ - assign it so that $\sum_{(u,w)\in E} \tilde{h}(u, w) = g(u)$ is satisfied.

It is also clear that this is the unique assignment to the edges in T which satisfies all these constraints and is consistent with the assignment to the edges of $E \setminus T$. This is because once the edges in the subtree of u has been fixed, there is a unique choice of $\tilde{h}(u, \text{parent}[u])$ that satisfies the parity constraint of u.

□

Note that if $\sum_{u\in V} g(u)$ is even, the procedure in the proof of Lemma 8.5 automatically ensures $\sum_{w\in N(v)} \tilde{h}(v, w) = g(v) \pmod 2$ is also satisfied. And if $\sum_{u\in V} g(u)$ is odd, the procedure automatically ensures $\sum_{w\in V} \tilde{h}(v, w) = g(v) \pmod 2$ is *not* satisfied. Therefore, we get the following corollary.

Corollary 8.6. *Let G be a connected undirected graph and let $g : V \to \mathbb{F}_2$ be any map.*

1. *If $\sum_{v\in V} g(v) = 1 \pmod 2$, then for any $v \in V$ there exists an assignment $\tilde{h} \in \mathbb{F}_2^E$ such that $\sum_{w\in N(u)} \tilde{h}(u, w) = g(u) \pmod 2$ for each $u \neq v$, and $\sum_{w\in N(v)} \tilde{h}(v, w) \neq g(v) \pmod 2$.*

2. *If $\sum_{v\in V} g(v) = 0 \pmod 2$, there exists an assignment $\tilde{h} \in \mathbb{F}_2^E$ such that $\sum_{w\in N(u)} \tilde{h}(u, w) = g(u) \pmod 2$ is satisfied for all $u \in V$.*

Now we define the following hard distribution for each $\rho \in P$, when $|\rho| \leq \frac{n}{2000}$.

8.1.2 The hard distrbution

Our goal in this subsection is to define for each $\rho \in P$ a distribution $\mu = \mu_\rho$ on the unfixed variables so that the requirement in Definition 8.1 is satisfied.

Definition 8.7. Let $\rho \in P$ be a valid partial assignment which fixes at most $\frac{n}{1000}$ edges. Define the following:

1. Let $S_\rho = \{e|\rho(e) \neq^*\}$ and $U_\rho = \{e|\rho(e) =^*\}$

2. Define f_ρ as before (i.e. for each $v \in V$, $f_\rho(v) = 1 + \sum_{(u,v)\in S_\rho} \rho(u, v)$ - the interpretation of $f_\rho(v)$ is that among the unfixed edges incident to v, the number of edges fixed to 1 must be $f_\rho(v) \pmod 2$ in order to satisfy the original parity constraint for v).

3. Let C_ρ be the unique connected component in $G_\rho = (V, E \setminus S_\rho)$ whose total f_ρ is odd.

Now we describe the procedure of sampling from μ. Note that the values of edges in S_ρ are fixed; we have to define a distribution on the unfixed edges. We do this as follows.

DTFooling

Input:

- Graph $G = (V, E)$
- A valid partial assignment $\rho \in \{0, 1, *\}^E$, $\rho \in P$

Output: A sample $z \in \{0, 1\}^E$ from μ_ρ

Sampling procedure

- For every vertex $v \in C_\rho$, fix an arbitrary spanning tree T_v of C rooted at v.
- Uniformly at random pick a vertex $v \in C_\rho$. Let $\tilde{z}$ be a uniformly random assignment to the unfixed edges in C not in T_v. Extend $\tilde{z}$ to the unique assignment z to all unfixed edges of C_ρ as guaranteed in Lemma 8.5 so that for all $u \neq v$, $f_\rho(u) = \sum_{(u,w) \in U_\rho} z(u, w)$ is satisfied.
- For every other connected component C', pick an arbitary spanning tree (with an arbitrary root). Give a uniformly random assignment to the non-tree edges; then fix the values of the tree edges according to Lemma 8.5. (Note that this assignment satisfies the parity constraint of all vertices in C' by the remark following Lemma 8.5.)

Let $A_{\rho,v}$ denote the set of all $z \in \mathbb{F}_2^E$ that are consistent with ρ and satisfy the following:

1. For all $u \neq v$, $\sum_{w \in N(u)} z(u, w) = f_\rho(u)$.
2. $\sum_{w \in N(v)} z(v, w) = 1 + f_\rho(v)$.

We make the following remark now.

Remark 8.1. *Let ρ be any valid (partial) assignment to edges of G. Then,*

1. *$A_{\rho,v}$ is an affine space in $\mathbb{F}_2^E$, for each $v \in C_\rho$.*
2. *DTFooling picks a random $v \in C_\rho$ and then samples a random point in $A_{\rho,v}$.*

For any $z \in \text{supp}(\mu)$, the parity constraint is violated for exactly one vertex (the vertex which was chosen as the root of the spanning tree of the odd connected component). Call this vertex *root*(z).

Before proving the hardness, we note down some properties of the distribution.

8.1.3 Conditional Distribution of the Root is Uniform

We prove a useful property of the distribution sampled by DTFooling, given a valid partial assignment ρ. The idea is when a decision tree queries bits from an assignment z to the edges sampled according to μ_ρ, the graph G_ρ starts splitting into further smaller components. The decision tree knows at every instant in which component root(z) lies,as there is always a unique odd component. The lemma below ensures that conditioned on what the decision tree has observed so far, the distribution of root(z) remains uniform over all vertices in the odd component.

In the following let $E_\rho \subseteq E$ be the set of edges free in ρ. For convenience of the reader, we re-state the definition of f_ρ in Definition 8.3 here.

- $f_\rho \in \mathbb{F}_2^V, f_\rho(v) = 1 + \sum_{(v,w)\in E\setminus E_\rho} \rho(v,w)$

Lemma 8.8. *Let ρ be a valid partial assignment and $\mu = \mu_\rho$ be the distribution in Definition 8.7. Let $S \subseteq E_\rho$ be a subset of free edges and $\alpha \in \mathbb{F}_2^S$ be an assignment to S. Define $f_\alpha(v) = f_\rho(v) + \sum_{(v,w)\in S} \alpha(v,w)$. Suppose the components of $(V, E_\rho \setminus S)$ are $C_1, C_2, \cdots, C_k$, where $\sum_{v\in C_1} f_\alpha(v) = 1$, and for all $j \neq 1$, $\sum_{v\in C_j} f_\alpha(v) = 0$. Then, the following are true:*

1. *For any $u \in C_1$, let $S_u := \{x \in \mathbb{F}_2^{E_\rho\setminus S} \mid root(x,\alpha) = u\}$. Then, $|S_u| = |S_v|$, for all $v \in C_1$.*
2. *The distribution of $root(z)$ as z is sampled from $\mu = \mu_\rho$ conditioned on $z|_S = \alpha$, is uniform on C_1.*

Proof. We shall prove point (2) (about the conditional distribution of the root being uniform on the odd component). En-route, we shall also end up proving point (1).

Conditioned on $z|_S = \alpha$, the root cannot lie in any of $C_2, C_3, \cdots, C_k$. (Reason: after we choose the root, only the parity constraint at the root is violated; other parity constraints are satisfied. But after fixing S to α, it is not possible to satisfy all parity constraints of C_1 simultaneously as the sum of the modified parity constraints of C_1 is odd.)

So we need to show that for all $u \in C_1$, $\Pr_\mu[\text{root}(z) = u | z_S = \alpha]$ is a non-zero quantity independent of u. Note that $C_1 \subseteq C_\rho$ (recall that C_ρ is the unique odd component of G_ρ; for a justification see Remark 8.2). Since for all $u \in C_1$, $\Pr_\mu[\text{root}(z) = u] = \frac{1}{|C_\rho|}$ is a non-zero quantity independent of u, by Bayes' rule it suffices to show that for all $u \in C_1$, $\Pr_\mu[z_S = \alpha | \text{root}(z) = u]$ is a non-zero quantity independent of u.

Let $M \in \mathbb{F}_2^{V\times E}$ be the edge-vertex incidence graph of G. Let $\gamma_v \in \mathbb{F}_2^V$ be the following vector:

$$\gamma_v(u) = \begin{cases} f_\rho(u) & \text{if } u \neq v \\ 1 + f_\rho(u) & \text{otherwise} \end{cases}$$

Once v is chosen as the root, the sampling procedure samples a uniformly random element of the affine space $\{z | Mz = \gamma_v\}$.

Let $S = \{r_1, r_2, \cdots, r_{|S|}\}$. Let $N \in F^{S\times E}$ be the matrix whose j-th row is the standard basis vector at coordinate r_j. Once u is chosen as the root, $z|_R = \alpha$ if and only if z satisfies the following equation:

$$\left[\frac{M}{N}\right] z = \left[\frac{\gamma_u}{\alpha}\right]$$

Let

$$J = \left[\frac{M}{N}\right]$$

Conditioned on satisfying $Mx = \gamma_u$, the probability of satisfying $z|_R = \alpha$ is either $2^{\text{rank}(M)-\text{rank}(J)}$ (if there is a solution) or 0 (if there is some inconsistency in the right hand sides of the system of equations). (Indeed, for $v \notin C_1$ there is an inconsistency in the right hand sides - as noted above.)

Now for all $v \in C_1$, we shall show there is a z such that $\text{root}(z) = v$ and $z|_S = \alpha$ – this will show that for $v \in C_1$, $\Pr[z|_S = \alpha | \text{root}(z) = v]$ is a non-zero quantity independent of v. To show this, we construct an assignment as follows:

- For each C_j, choose a spanning tree Q_j disjoint from S.
- Assign $z|_S = \alpha$, and for all non-tree edges not in S, give an arbitrary assignment.
- Assign the values of the edges in Q_1 according to Lemma 8.5 with v as root.
- For $j > 1$, set the edges of Q_j according to Lemma 8.5 (with any arbitrary root).

This assignment satisfies $z|_S = \alpha$ by construction; moreover, it also satisfies $Mz = \gamma_v$ because of Lemma 8.5 and the remark following it (Corollary 8.6). This argument, at one go, establishes both properties (1) and (2) claimed in our lemma. $\square$

8.1.4 Proving Hardness

Now we prove $(p, q)-$hardness for ordinary decision trees.

Theorem 8.9. *Let $G = (V, E)$ be a $(|V|, d, 1/3000)$-spectral expander with $|V|$ odd. Let P be the set of partial assignments defined in Definition 8.3. Let $\rho \in P$ be a partial assignment with $|\rho| = p \leq \frac{n}{2000}$. Let μ be the distribution defined in Definition 8.7*[4]*. Let T be any decision tree making at most $q \leq \frac{n}{2000}$ queries. Sample $z \leftarrow \mu$. Then, with probability $\geq 1/3$, the partial assignment seen by the tree after q queries (this includes the edges fixed by ρ and the edges queried by T) also lies in P.*

Proof. We fix some notation that will be used in the rest of the proof.

1. At time-step j, the partial assignment seen by the tree is ρ_j (this includes the edges fixed by ρ and the edges queried by T).
2. The set of edges fixed by ρ_j is E_j. The corresponding graph is $G_j = (V, E \setminus E_j)$.
3. Define $f_j(v) = 1 + \sum_{(v,w) \in E_j} \rho_j(v, w)$.
4. Let the connected components of G_j be $D_1, D_2, \cdots, D_k$. Because of the way μ is defined, there always exists exactly one connected component D_l such that $\sum_{v \in C_l} f_j(v) = 1$ (for a justification, see remark 8.2). Call this *the odd component of G_j*. Denote it by C_j.

Before proceeding, we make some remarks.

Remark 8.2. *After some vertex v is chosen to be the root, it is guaranteed that the parity constraint of all vertices other than v is satisfied. It is also known that all parity constraints are not satisfiable simultaneously (since sum of the right hand sides is odd). So, after we know the value of some edges (say given by the partial assignment σ), after removing those edges, exactly one connected component has odd $\sum f_\sigma$ - and the root lies in this component. This means that after z is sampled according to μ, condition (i) defining membership in P is always satisfied. Only condition (ii) (which stipulates that the odd component must have large size) can possibly be violated*

Remark 8.3. *Suppose after querying an edge, in G_{j+1} the component C_j splits into $C_j = A \cup B$. Initially $\sum_{v \in C_j} f_j(v)$ is odd. After querying the edge, exactly one of $\sum_{v \in A} f_{j+1}(v)$ and $\sum_{v \in B} f_{j+1}(v)$ is odd - and the root must lie in the component where the sum is odd.. Suppose $\sum_{v \in A} f_{j+1}(v)$ is*

[4] μ is a distribution on $\mathbb{F}_2^n$ such that every $z \in \text{supp}(\mu)$ is consistent with ρ

odd. Then, the decision tree now knows the root lies in $A = C_{j+1} \subseteq C_j$. Thus, the decision tree has made some progress in determining the location of the root.

We want to say that the decision tree can never make too much progress - our tool here is Lemma 8.8, which says that the decision tree does not know anything about the root other than it lies in the current odd component.

We start with a crucial lemma.

Lemma 8.10. *At any time-step j, the largest connected component of G_j must have size $\geq n\left(1 - \frac{1}{50d}\right)$*

Proof. Suppose not; let time-step j be a time step where all the connected components of G_j have size $< n\left(1 - \frac{1}{50d}\right)$. We then greedily pick a subset of the connected components whose union T has size in the interval $\left[\frac{n}{100d}, n - \frac{n}{100d}\right]$. Cheeger's inequality (Lemma 6.1) then implies the cut $E(T, V \setminus T)$ has at least $\frac{n}{500}$ edges.

This means the current partial assignment fixes at least $n/500$ edges. However, the current partial assignment can only fix $p + q \leq n/1000$ edges. $\square$

Now, every time the decision tree queries an edge, we make it pay us some coins as follows. Suppose the current partial assignment lies in P; the current graph is G_j and the current odd component is C_j, and the tree queries the edge e.

- If removing e keeps C_j connected, the tree does not have to pay anything.
- Suppose removing e splits C_j into two components: $C_j = A \cup B$. The value of e is revealed - and it determines in which component of A, B the root belongs to. Suppose the root lies in A. If $|A| \leq n/2$, the decision tree does not pay anything and wins the game. Otherwise, the decision tree has to pay $|B|$ coins.

(In other words: if, at any point of time, the largest component in G_j isn't the odd component, the decision tree wins the game. Otherwise, if the decision tree shrinks the size of the largest component by s, it must pay s coins.)

By Lemma 8.10, the decision tree only pays $\leq \frac{n}{50d}$ coins. So we start by awarding the decision tree a budget of $b = \frac{n}{50d}$ coins, and argue (by induction on number of coins remaining) that the decision tree loses the game with high probability. (The decision tree loses the game when it has to pay some coins but it is broke.)

At this point, we allow the decision tree to make as many as queries as it wants – as long as it maintains that the largest component has size $\geq n\left(1 - \frac{1}{50d}\right)$ (and therefore it does not use more than b coins). We prove the following statement by inducting on number of coins remaining.

Lemma 8.11. *Suppose the decision tree has c coins remaining and has not won the game yet. Then, the probability it wins the game is $\leq \frac{3c}{n} + \frac{1}{10}$.*

Proof. We induct on c. Consider the base case $c = 0$: the decision tree has no coins remaining. Let the current odd component be C. The first time the tree splits C, the root must lie in the smaller component for the decision tree to win. Suppose the tree queries an edge e and C splits into $C = A \cup B$ where $|A| \geq n\left(1 - \frac{1}{50d}\right)$. Before querying e, the conditional distribution of the root was uniform on C by Lemma 8.8. Conditioned on the partial assignment revealed before querying e, the probability the root lies in A is $|A|/|C|$. The probability the tree wins the game is thus

$$\frac{|A|}{|C|} \leq \frac{\frac{n}{50d}}{n\left(1 - \frac{1}{50d}\right)} \leq \frac{1}{10},$$

so the base case is true.
Now we handle the inductive step. Suppose the tree has c coins. Suppose the current odd component is C_j, with $|C_j| \geq n\left(1 - \frac{1}{50d}\right)$. Suppose the decision tree queries e and removing e splits C_j into $C_j = A \cup B$, where A is the larger component $\left(|A| \geq n\left(1 - \frac{1}{50d}\right)\right)$. The tree wins this game at this stage if the root lies in B, otherwise it pays $|B|$ coins and proceeds to the next stage. Before querying e the distribution of the root is uniform on C, so the probability it lies in B is $\frac{|B|}{|A|+|B|} \leq \frac{3|B|}{n}$. If the root does not lie in B, the decision tree has $c - |B|$ coins remaining, so then it can win the came with probability at most $\frac{3(c-|B|)}{n} + \frac{1}{10}$ by the inductive hypothesis. By union bound, the probability the tree wins the game is at most

$$\frac{3|B|}{n} + \frac{3(c-|B|)}{n} + \frac{1}{10} = \frac{3c}{n} + \frac{1}{10}$$

□

Since the decision tree starts off with $\frac{n}{50d}$ coins, it can win with probability at most $\frac{3}{50d} + \frac{1}{10} < \frac{2}{3}$. Hence, with probability $\geq \frac{1}{3}$, the partial assignment seen by the decision tree after q queries lies in P. □

Now we have all the ingredients require to prove $(n/2000, n/2000)$-DT-hardness of Φ.

Proof. *(of Theorem 8.2)* Choose the set of partial assignments P as defined in Section 8.1.1. We have already established that this set P satisfies all requirements in Definition 8.1 defining DT-hardness.

- **No falsification** and **Downward Closure**: Established in Lemma 8.4.
- **Hardness against decision trees:** Established in Theorem 8.9.

This completes the proof. □

8.2 Onto (p, q)-PDT Hardness

Theorem 8.9 shows that Tseitin contradiction over a constant degree expander is $(\Omega(n), \Omega(n))$-DT hard. In this section, we show how to modify the proof of Theorem 8.9 (in a white-box manner) to prove the original requirement of (p, q)-PDT hardness for $p, q = \Omega(n)$.

Recall the condition for (p, q)-PDT hardness:

- **Hardness:** Let $A \subseteq \mathbb{F}_2^{nb}$ be an affine space with $|\hat{\mathrm{Cl}}(A)| \leq p$. Let $y \in \mathbb{F}_2^{\mathrm{VARS}(\mathrm{Cl}(A))}$ be an extendable closure assignment for A such that $\alpha = G(y) \in P$. Then, there exists a distribution $\mu = \mu(\alpha)$ on $\mathbb{F}_2^n$ such that the following properties hold:

1. $z|_{\mathrm{Cl}(A)} = \alpha$ for all $z \in \mathrm{supp}(\mu)$
2. Let T be any parity decision tree (on nb bits of input) of depth at most q. For any x, define $\tilde{A}(x) = A_T(x) \cap A \cap C_y$. When x is sampled according to $G^{-1}(\mu) \cap C_y$, with probability $\geq 1/3$, it holds that $G(x|_{\mathrm{Cl}(\tilde{A}(x))}) \in P$.

We state the main result of this section below.

Theorem 8.12. *The set of valid assignments P, as specified in Definition 8.3, and the hard distributions μ_ρ sampled by DTFooling for each $\rho \in P$, together exhibit that Tseitin contradictions on graphs $G = (V, E)$ (with $|V|$ odd) which is are sufficiently good expanders, are $(\Omega(n), \Omega(n))$-PDT hard, where n is the number of variables in the Tseitin CNF (which is also equal to the number of edges in G).*

For a PDT T and an input x, denote by $T_j(x)$ the following affine space:

- Let $l_1, l_2, \cdots, l_j$ be the first j linear forms queried by T on input x, and let $c_1, c_2, \cdots, c_j$ be the responses. Then, $T_j(x) = \{\tilde{x} | \langle l_i, \tilde{x} \rangle = c_i \forall i \in [j]\}$ (in case j is more than the number of queries issued by T on x, define $T_j(x) = T_k(x)$ where k is the number of queries issued by T on x).

Let $\tilde{A}_j(x) = T_j(x) \cap A \cap C_y$, for any PDT T, and any affine space A. Also, define $\mathcal{A}(x) = \tilde{A}_r(x)$ where r is the number of queries made by T on input x.

Since the amortized closure of $\tilde{A}_0(x) := A \cap C_y$ is at most $|\hat{\mathrm{Cl}}(A)| \leq p$, we know by Lemma 3.27 that $|\mathrm{Cl}(\mathcal{A}(x))|$ is at most $p + q$. We construct a modified PDT $\tilde{T}$ such that it makes some additional queries, but it still holds that $|\mathrm{Cl}(\mathcal{A}(x))| \leq p + q$ (by Lemma 3.27). We will prove the lower bound for the modified PDT $\tilde{T}$, which also implies the lower bound for the original PDT.
The modification is as follows: suppose originally T was about to query a linear form ℓ, and that would add k new blocks $B_1, B_2, \cdots, B_k$ to $\mathrm{Cl}(\tilde{A})$. Then, the modified PDT $\tilde{T}$ first queries all bits of $B_1, B_2, \cdots, B_k$, and then it queries ℓ. (After this stage, $\mathrm{Cl}(\tilde{A})$ is the same for both T and $\tilde{T}$).
So we can describe the behaviour of $\tilde{T}$ as follows:

- At any time-step j, $\tilde{T}$ has queried all bits of a subset of blocks $S_j \subset [n]$. At this point, it holds that $\mathrm{Cl}(\tilde{A}_j) = S_j$.
- At each time-step, $\tilde{T}$ can either query all bits of some new blocks, or query a general linear form ℓ such that querying ℓ does not change $\mathrm{Cl}(\tilde{A})$.
- In total, it can query at most $p + q$ blocks.

The crucial observation is the following: suppose currently $\tilde{T}$ has so far queried the subset S_j of blocks and the current closure assignment is $y_j \in \mathbb{F}_2^{\mathrm{VARS}(S_j)}$. Suppose $\alpha_j = G(y_j) \in P$. Then, the distribution of $\mathrm{root}(G(x))$ as x is sampled from $G^{-1}(\mu) \cap \tilde{A}_j \cap C_{y_j}$ [5], is $o(n^{-20})$-close to uniform (in statistical distance) over the unique odd component of α_j. This fact is formally stated in the next lemma below, which is the lifted analogue of Lemma 8.8. Indeed, the proof of Lemma 8.13 below will make use of Lemma 8.8.

Recall that for a partial assignment $\alpha \in \mathbb{F}_2^E$, the graph G_α is defined to be the original graph G with the edges fixed by α removed.

Lemma 8.13. *Let $A \subseteq \mathbb{F}_2^{nb}$ be an affine subspace and let $y_1 \in F^{VARS(Cl(A))}$ be an extendable closure assignment such that $\alpha = G(y_1) \in P$. Let $\mu = \mu_\alpha$ be the distribution in definition 8.7* [6].

[5] Note that since $\tilde{T}$ queries all bits of $\mathrm{Cl}(A_j)$ separately, we have that $\tilde{A}_j \subseteq C_{y_j}$. But we still write $G^{-1}(\mu) \cap \tilde{A}_j \cap C_{y_j}$ in the subscript for notational consistency.

[6] μ is a distribution on $\mathbb{F}_2^n$ such that every point in $\mathrm{supp}(\mu)$ is consistent with $G(y)$. We emphasize that μ only depends on y, not on A_2 or y_2

Let $A_2 \subseteq A$ be an affine space, and let $y_2 \in \mathbb{F}_2^{VARS(Cl(A_2))}$ be an extendable closure assignment for A_2 such that $\beta = G(y_2) \in P$ and y_2 extends y_1. Let C be the unique odd component of G_β. Then, the conditional distribution of root$(G(x))$, as x is sampled from $G^{-1}(\mu) \cap C_{y_2} \cap A_2$ is $o(n^{-20})$-close to UNIFORM(C) in ℓ_1 distance.

Proof. As before, in the proof of Lemma 8.8, conditioned on $G(y_2)$, the root cannot lie in any of the even components.

Claim 8.14. For all $u, v \in C$,

$$\frac{\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = u]}{\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = v]} \in \left[\frac{1-o(n^{-80})}{1+o(n^{-80})}, \frac{1+o(n^{-80})}{1-o(n^{-80})}\right]$$

Proof. Let S_u be the set of $\tilde{z} \in \mathbb{F}_2^{[n]\setminus \mathrm{Cl}(A_2)}$ such that $(z,\beta) \in \mathrm{supp}(\mu)$ and $\mathrm{root}(z,\beta) = u$. Define S_v similarly. Note that, $|S_u| = |S_v|$ by Property (1) in Lemma 8.8.
Let $\tilde{\mu}$ be projection/marginal of μ on $\mathbb{F}_2^{[n]\setminus \mathrm{Cl}(A_2)}$ We re-write the numerator on the LHS in our claim as follows:

$$\begin{aligned}\Pr_{x\sim G^{-1}(\mu)\cap C_{y_2}\cap A_2}[\mathrm{root}(G(x)) = u] &= \sum_{\tilde{z}\in S_u} \Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[G(x) = (\tilde{z},\beta)] \\ &= \sum_{\tilde{z}\in S_u} \frac{\Pr_{x\sim G^{-1}(\mu)}[(x \in A_2 \cap C_{y_2}) \wedge (G(x) = (\tilde{z},\beta))]}{\Pr_{x\sim G^{-1}(\mu)}[x \in A_2 \cap C_{y_2}]}\end{aligned} \tag{8.1}$$

Re-writing things, letting $M = |\mathrm{supp}(\mu)|$, we get the following:

$$\begin{aligned}&\Pr_{x\sim G^{-1}(\mu)}[(x \in A_2 \cap C_{y_2}) \wedge (G(x) = (\tilde{z},\beta))] \\ &\qquad = \Pr_{x\sim G^{-1}(\mu)}[G(x) = (\tilde{z},\beta)] \cdot \Pr_{x\sim G^{-1}(\tilde{z},\beta)}[x \in (A_2 \cap C_{y_2})] \\ &\qquad = \frac{\Pr_{x\sim G^{-1}(\tilde{z},\beta)}[x \in (A_2 \cap C_{y_2})]}{M}\end{aligned} \tag{8.2}$$

Note that in the second equality, we have simply used the fact that μ samples every point in its support uniformly at random. Let, $\tilde{x} = x|_{[nb]-\mathrm{Cl}(A_2)}$ and $\hat{x} = x|_{\mathrm{Cl}(A_2)}$ and $p = \Pr_{\hat{x}\sim G^{-1}(\beta)}[\hat{x} = y_2]$. Finally, let $(A_2)_{y_2}$ denote the affine subspace of $\mathbb{F}_2^{[nb]-b\mathrm{Cl}(A_2)}$ induced as follows:

$$(A_2)_{y_2} := \{w \in \mathbb{F}_2^{[nb]-b\mathrm{Cl}(A_2)} \mid (w, y_2) \in A_2\}.$$

Then, continuing from (8.2), we get the following:

$$\begin{aligned}\Pr_{x\sim G^{-1}(\mu)}[(x \in A_2 \cap C_{y_2}) \wedge (G(x) = (\tilde{z},\beta))] &= \frac{\Pr_{\tilde{x}\sim G^{-1}(\tilde{z})}[\tilde{x} \in (A_2)_{y_2}] \cdot \Pr_{\hat{x}\sim G^{-1}(\beta)}[\hat{x} = y_2]}{M} \\ &= \frac{p}{M}\Pr_{\tilde{x}\sim G^{-1}(\tilde{z})}[\tilde{x} \in (A_2)_{y_2}] \\ &= \frac{p}{M}\frac{|G^{-1}(\tilde{z}) \cap (A_2)_{y_2}|}{|G^{-1}(\tilde{z})|}\end{aligned} \tag{8.3}$$

Combining (8.1) and (8.3), we get

$$\Pr_{x\sim G^{-1}(\mu)\cap C_{y_2}\cap A_2}[\mathrm{root}(G(x)) = u] = \frac{p}{M \cdot \Pr_{x\sim G^{-1}(\mu)}[x \in A_2 \cap C_{y_2}]} \sum_{\tilde{z}\in S_u} \frac{|G^{-1}(\tilde{z}) \cap (A_2)_{y_2}|}{|G^{-1}(\tilde{z})|}$$

Let $d = \mathrm{codim}(A_2)_{y_2}$. Note that $(A_2)_{y_2}$ is a nice affine space, so by Lemma 5.7 we have for all $\tilde{z}$,

$$\frac{|G^{-1}(\tilde{z}) \cap (A_2)_{y_2}|}{|G^{-1}(\tilde{z})|} \in \left[\frac{1-o(n^{-90})}{1+o(n^{-90})} \times \frac{1}{2^d}, \frac{1+o(n^{-90})}{1-o(n^{-90})} \times \frac{1}{2^d}\right]$$

Thus,

$$\Pr_{x\sim G^{-1}(\mu)\cap C_{y_2}\cap A_2}[\mathrm{root}(G(x)) = u] \in \frac{p|S_u|}{\Pr_{x\sim G^{-1}(\mu)}[x \in A_2 \cap C_{y_2}]M2^d}\left[\frac{1-o(n^{-90})}{1+o(n^{-90})}, \frac{1+o(n^{-90})}{1-o(n^{90})}\right]$$

The same bounds hold for the denominator (the corresponding expression for v), so we have that (since, by Property (1) of Lemma 8.8, $|S_u| = |S_v|$)

$$\frac{\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = u]}{\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = v]} \in \left[\frac{1-o(n^{-80})}{1+o(n^{-80})}, \frac{1+o(n^{-80})}{1-o(n^{-80})}\right]$$

□

The proof of Lemma 8.13 follows easily from Claim 8.14. We shall show that for all $u \in C$,

$$\left|\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = u] - \frac{1}{|C|}\right| \le o(n^{-70})$$

This means $\mathrm{root}(G(x))$ and UNIFORM(C) are $o(n^{-70})$ close in ℓ_∞ distance, which implies they are $o(n^{-50})$ close in ℓ_1 distance.

There exists a $v \in C$ such that $\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = v] \ge \frac{1}{|C|}$ (since the support of $\mathrm{root}(G(x))$ conditioned on $x \in A_2 \cap C_{y_2}$ is C), so

$$\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = u] \ge \frac{1}{|C|} \times \frac{1-o(n^{-80})}{1+o(n^{-80})} \ge \frac{1}{|C|} - o(n^{-70})$$

Similarly

$$\Pr_{x\sim G^{-1}(\mu)\cap A_2\cap C_{y_2}}[\mathrm{root}(G(x)) = u] \le \frac{1}{|C|} + o(n^{-70})$$

The result follows.

□

We have all the technical results in place for proving Theorem 8.12, the main result of this section stating that Tseitin contradictions on expanding graphs are $\Omega(n), \Omega(n)$)-PDT hard.

Proof Sketch of Theorem 8.12. First modify the given PDT as described above (after the statement of Theorem 8.12). The rest of the proof proceeds identically to the proof of the (p, q)-DT hardness where Lemma 8.11 made repeated use of the (perfectly) uniform distribution of the root as given by Lemma 8.8. Here we have to straightforwardly modify the argument in the proof of Lemma 8.11 by invoking Lemma 8.13 instead of Lemma 8.8 – every time the PDT queries some new blocks, it has to pay some money depending on how much the odd connected component shrinks. Since the conditional distribution of the root is close to uniform on the odd component, with high probability it will end up in the largest component during each cut (in the same way as in the unlifted case). The linear queries which don't affect the closure do not affect the conditional distribution of the root in any significant way. □

For completeness, we have presented a complete self-contained proof of Theorem 8.12, using Lemma 8.13, in Appendix B

9 Putting everything together

With Theorem 8.12 in hand, we are now in a position to prove our main result, Theorem 1.1.

Proof. (of Theorem 1.1) Let Φ be the Tseitin contradiction on G. Theorem 8.12 shows that that Tseitin contradiction over an expander is $(\Omega(n), \Omega(n))$-hard. It is also known that the Inner Product gadget has exponentially small Fourier coefficients (Theorem 6.2). Applying Theorem 7.2, we get the following result:

- Any size s refutation of $\Phi \circ \mathrm{IP}$ must require depth $\Omega\left(\frac{n^2}{\log(s)}\right)$

Note that the number of variables in $\Phi \circ \mathrm{IP}$ is $N = O(n \log(n))$. We can also interpret the result as follows:

- Any depth $N^{2-\epsilon}$ $\mathrm{Res}(\oplus)$ refutation of $\Phi \circ \mathrm{IP}$ requires size $\exp(\tilde{\Omega}(N^{\epsilon}))$

This is what we wanted to show. □

References

[ABRW04] 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.

[AI25] Yaroslav Alekseev and Dmitry Itsykson. Lifting to bounded-depth and regular resolutions over parities via games. In Michal Koucký and Nikhil Bansal, editors, *Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC 2025, Prague, Czechia, June 23-27, 2025*, pages 584–595. ACM, 2025.

[BCD24] Sreejata Kishor Bhattacharya, Arkadev Chattopadhyay, and Pavel Dvorák. Exponential separation between powers of regular and general resolution over parities. In Rahul Santhanam, editor, *39th Computational Complexity Conference, CCC 2024, July 22-25, 2024, Ann Arbor, MI, USA*, volume 300 of *LIPIcs*, pages 23:1–23:32. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.

[BIK+92] Paul Beame, Russell Impagliazzo, Jan Krajícek, Toniann Pitassi, Pavel Pudlák, and Alan R. Woods. Exponential lower bounds for the pigeonhole principle. In S. Rao Kosaraju, Mike Fellows, Avi Wigderson, and John A. Ellis, editors, *Proceedings of the 24th Annual ACM Symposium on Theory of Computing, May 4-6, 1992, Victoria, British Columbia, Canada*, pages 200–220. ACM, 1992.

[BK23] Paul Beame and Sajin Koroth. On disperser/lifting properties of the index and inner-product functions. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPIcs*, pages 14:1–14:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023.

[BW01] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow - resolution made simple. *J. ACM*, 48(2):149–169, 2001.

[CD25] Arkadev Chattopadhyay and Pavel Dvorák. Super-critical trade-offs in resolution over parities via lifting. In Srikanth Srinivasan, editor, *40th Computational Complexity Conference, CCC 2025, August 5-8, 2025, Toronto, Canada*, volume 339 of *LIPIcs*, pages 24:1–24:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025.

[CDK$^+$17] Arkadev Chattopadhyay, Pavel Dvorák, Michal Koucký, Bruno Loff, and Sagnik Mukhopadhyay. Lower bounds for elimination via weak regularity. In Heribert Vollmer and Brigitte Vallée, editors, *34th Symposium on Theoretical Aspects of Computer Science, STACS 2017, March 8-11, 2017, Hannover, Germany*, volume 66 of *LIPIcs*, pages 21:1–21:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[CFK$^+$21] Arkadev Chattopadhyay, Yuval Filmus, Sajin Koroth, Or Meir, and Toniann Pitassi. Query-to-communication lifting using low-discrepancy gadgets. *SIAM J. Comput.*, 50(1):171–210, 2021.

[Che71] Jeff Cheeger. *A Lower Bound for the Smallest Eigenvalue of the Laplacian*, pages 195–200. Princeton University Press, Princeton, 1971.

[CMSS23] Arkadev Chattopadhyay, Nikhil S. Mande, Swagato Sanyal, and Suhail Sherif. Lifting to parity decision trees via stifling. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPIcs*, pages 33:1–33:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023.

[EGI24] Klim Efremenko, Michal Garlík, and Dmitry Itsykson. Lower bounds for regular resolution over parities. 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 640–651. ACM, 2024.

[EI25] Klim Efremenko and Dmitry Itsykson. Amortized closure and its applications in lifting for resolution over parities. In Srikanth Srinivasan, editor, *40th Computational Complexity Conference, CCC 2025, August 5-8, 2025, Toronto, Canada*, volume 339 of *LIPIcs*, pages 8:1–8:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025.

[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.

[GPT22] Svyatoslav Gryaznov, Pavel Pudlák, and Navid Talebanfard. Linear branching programs and directional affine extractors. In Shachar Lovett, editor, *37th Computational Complexity Conference, CCC 2022, July 20-23, 2022, Philadelphia, PA, USA*, volume 234 of *LIPIcs*, pages 4:1–4:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[Hak85] Armin Haken. The intractability of resolution. *Theor. Comput. Sci.*, 39:297–308, 1985.

[IS14] Dmitry Itsykson and Dmitry Sokolov. Lower bounds for splittings by linear combinations. In Erzsébet Csuhaj-Varjú, Martin Dietzfelbinger, and Zoltán Ésik, editors, *Mathematical Foundations of Computer Science 2014 - 39th International Symposium, MFCS 2014, Budapest, Hungary, August 25-29, 2014. Proceedings, Part II*, volume 8635 of *Lecture Notes in Computer Science*, pages 372–383. Springer, 2014.

[LPS88] A Lubotzky, R Phillips, and P Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, September 1988.

[Mar73] G.A. Margulis. Explicit constructions of expanders. *Problemy Pereda ˇci Informacii, 9(4):71–80,*, 1973.

[MP97] Alexis Maciel and Toniann Pitassi. On $\mathrm{ACC}^0[p^k]$ Frege proofs. In Frank Thomson Leighton and Peter W. Shor, editors, *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, El Paso, Texas, USA, May 4-6, 1997*, pages 720–729. ACM, 1997.

[Urq87] Alasdair Urquhart. Hard examples for resolution. *J. ACM*, 34(1):209–219, 1987.

Appendices

A Facts about Amortized Closure

Proof. **(of Lemma 3.26)** $\mathrm{Cl}(V) = \mathrm{Cl}(W)$ follows from the definition of closure.

To see that $\hat{\mathrm{Cl}}(V) = \hat{\mathrm{Cl}}(W)$, consider the following set: $V_{in} = \{v | v \in \mathrm{span}(V), \mathrm{supp}(v) \subseteq \mathrm{Cl}(V)\}$. Let $B_{in} = \{v_1, v_2, \cdots, v_t\}$ be a basis for V_{in}, and extend it to a basis $B = \{v_1, v_2, \cdots, v_t, w_1, w_2, \cdots, w_k\}$ for $\mathrm{span}(V)$. Now replace V in the statement by B - since $\hat{\mathrm{Cl}}$ depends only on the span, $\hat{\mathrm{Cl}}(V)$ and $\hat{\mathrm{Cl}}(W)$ do not change.
Define the following matrices:

$$M = \begin{bmatrix} v_1 \\ \hline v_2 \\ \hline \cdots \\ \hline v_t \end{bmatrix}$$

$$\mathcal{I} = \left[e_{b,j} |_{b \in \mathrm{Cl(V)}, j \in [b]} \right] \in \mathbb{F}_2^{b|\mathrm{Cl}(V)| \times nb}$$

$$\mathcal{K} = \begin{bmatrix} M \\ \hline w_1 \\ \hline \cdots \\ \hline w_t \end{bmatrix}$$

$$\mathcal{J} = \begin{bmatrix} \mathcal{K} \\ \hline \mathcal{I} \end{bmatrix}$$

In words, the matrix $\tilde{I}$ stacks the standard basis vectors $e_{j,k}$ for each $j \in \mathrm{Cl}(V), k \in [b]$ together. The matrix $\mathcal{K}$ is obtained by stacking the vectors in V. The matrix $\mathcal{J}$ is obtained by stacking the vectors in W.

By the previous fact, the set $\mathrm{Cl}(V)$ is acceptable for the set of vectors V_{in}.
Let $\mathrm{Cl}(V) = \{s_1, s_2, \cdots, s_t\}$, and suppose we choose indices $\{c_1, c_2, \cdots, c_t\}$ such that c_j lies in block s_j and the corresponding columns of M are linearly independent. Consider any maximally acceptable set for w, $\tilde{S}$. Note that $\tilde{S}$ contains $\mathrm{Cl}(V)$. Since $\tilde{S}$ is acceptable, for each block $\tilde{p} \in \tilde{S}$ we can pick an index $\mathrm{IND}(\tilde{p})$ in the $\tilde{p}$-th block such that the corresponding columns in $\mathcal{J}$ are linearly independent. Given any such choice of indices, for each $s_j \in \mathrm{Cl}(V)$, we can change $\mathrm{IND}(s_j)$ to c_j and the new choice will still be valid. (Reason: the corresponding columns in c_j are linearly independent in M itself; and the other columns are 0 within the first t rows.)
Now it is clear that any maximal acceptable set for $\mathcal{J}$ is also acceptable for $\mathcal{K}$ - simply modify the IND function as mentioned above.

□

Proof. **(of Corollary 3.28)** By Lemma 3.27 we have that $\mathrm{Cl}(A) = \mathrm{Cl}(B)$, so A_y, B_y are nice affine spaces. It remains to show that $\mathrm{codim}(B_y) = \mathrm{codim}(A_y) + 1$.

Let $A = \{x|Mx = b\}$ and let the set of rows of M be $v_1, v_2, \cdots, v_k$. Let $B = \{x|\tilde{M}x = \tilde{b}\}$ where $\tilde{M}$ has $k+1$ rows, the first k of which are $v_1, v_2, \cdots, v_k$. Let the last row be w.

Let $S = \mathrm{Cl}(A) = \mathrm{Cl}(B)$. The set of defining linear forms of A_y is $v_1[\backslash S], v_2[\backslash S], \cdots, v_k[\backslash S]$ and the set of defining linear forms of B_y is $v_1[\backslash S], \cdots, v_k[\backslash S], w[\backslash S]$. We wish to show $w[\backslash S]$ is linearly independent from $v_1[\backslash S], v_2[\backslash S], \cdots, v_k[\backslash S]$. This is equivalent to showing that w does not lie in $\mathrm{span}(\{v_1, v_2, \cdots, v_k\} \cup \{e_{i,j}|i \in S, j \in [b]\})$. FTSOC assume w lies in $\mathrm{span}(\{v_1, v_2, \cdots, v_k\} \cup \{e_{i,j}|i \in S, j \in [b]\})$. Thus, $w = r + s$ for some $r \in \mathrm{span}(v_1, v_2, \cdots, v_k)$ and $s \in \mathbb{F}_2^n$ such that $\mathrm{supp}(s) \subseteq S$. Since Cl and $\hat{\mathrm{Cl}}$ of a set depend only on its linear span, we can WLOG replace w by s. Hence, assume $\mathrm{supp}(w) \subseteq S$.

By Lemma 3.26, we have that

$$\hat{\mathrm{Cl}}(\{v_1, v_2, \cdots, v_k, w\}) \subseteq \hat{\mathrm{Cl}}(\{v_1, v_2, \cdots, v_k\} \cup \{e_{i,j}|i \in S, j \in [b]\}) = \hat{\mathrm{Cl}}(\{v_1, v_2, \cdots, v_k\})$$

This is a contradiction, since we assumed that

$$|\hat{\mathrm{Cl}}(\{v_1, v_2, \cdots, v_k, w)| = |\hat{\mathrm{Cl}}(v_1, v_2, \cdots, v_k)| + 1$$

□

B Proof of $(\Omega(n), \Omega(n))$-PDT-hardness of Tseitin contradiction over an expander

In this Appendix we shall provide a self-contained proof that when Φ is the Tseitin contradiction over a sufficiently good expander, Φ is $\left(\frac{n}{2000}, \frac{n}{2000}\right)$ PDT-hard (which was Lemma 8.12 in the paper). As mentioned in section 8, the following proof simply combines the analogous unlifted proof of Theorem 8.7 with the equidistribution lemma as applied in Lemma 8.13.

Proof. **(of Lemma 8.12)**
We choose the set $P \subseteq \{0, 1, *\}^n$ to be the same as that in the proof of Lemma 8.7 (there exists exactly one component with sum $\sum f_\rho$ odd and that component has size at least $n/2$). We shall show this set of partial assignments is $(\Omega(n), \Omega(n))$ hard as defined in Definition 7.1.

We are given an affine space A with $\hat{\mathrm{Cl}}(A) \leq p = n/2000$ and a closure assignment $y \in \mathbb{F}_2^{\mathrm{Cl}(A)}$ such that $G(y) \in P$:. We start by re-introducing some notation for the convenience of the reader.

- For a PDT T, define $\mathcal{A}(T, x, j)$ to be the following affine space:
  - Let $\ell_1, \ell_2, \cdots, \ell_j$ be the first j queries made by T on x and let $b_1, b_2, \cdots, b_j$ be the responses. Define $\mathcal{A}_j(T, x) = \{x|\langle \ell_i, x\rangle = b_i \ \forall i \in [j]\}$. If T issues fewer than j queries on x, define $\mathcal{A}_j(T, x) = \mathcal{A}_q(T, x)$ where q is the number of queries issued by T on x.
  - Define $\tilde{\mathcal{A}}_j(T, x) = \mathcal{A}_j(T, x) \cap \mathcal{A} \cap C_y$.

We choose the distribution $\mu = \mu_{G(y)}$ to be the same distribution defined in section 8 (**DTFooling**). Following the definition of (p, q)-hardness in Definition 7.1, we have to show the following:

- Let T be a PDT of depth $\leq n/1000$ (working on nb bits). With probability $\geq 1/3$, as we sample $x \leftarrow G^{-1}(\mu) \cap A \cap C_y$, letting $B = \tilde{\mathcal{A}}(T, x, n/2000)$, we have that $G(x|_{\mathrm{Cl}(B)}) \in P$

(The other two requirements, **no falsification** and **downward closure** were already established in Lemma 8.4.)

Since $|\hat{\mathrm{Cl}}(A)| < n/2000$, by Lemma 3.27 we know that $|\hat{\mathrm{Cl}}(\mathcal{A}(T, x, n/2000))| \leq \frac{n}{1000}$ and therefore, $|\mathrm{Cl}(\mathcal{A}(T, x, n/2000))| \leq \frac{n}{1000}$.

We construct a modified PDT T' which makes a few more queries than T, but it still holds that $|\mathrm{Cl}(\mathcal{A}(T', x, n/2000))| \leq \frac{n}{1000}$. The modified PDT functions at follows:

- The PDT T' queries in time-steps. At time-step j, it simulates the j-th query of T and makes some additional queries. Denote by $\tilde{\mathcal{A}'}_j(x)$ to be the affine space after the j-th time-step – i.e., if K is the total number of queries issued by T' till time-step j, $\tilde{\mathcal{A}}'_j(x) = \tilde{\mathcal{A}}_K(T', x)$

The following invariant is maintained by T':

- At the end of time-step j, let $S_j = \mathrm{Cl}(\tilde{\mathcal{A}}_j)$ be the closure of the current affine space. Then, at this point, T' has queried all bits of $\mathrm{VARS}(S_j)$. It also holds that $S_j = \mathrm{Cl}(\tilde{\mathcal{A}}(T, x, j))$ (where recall that $\tilde{\mathcal{A}}(T, x, j)$ was the subspace of the original PDT T).

The simulation of T' happens as follows:

- Suppose the invariant is maintained till time-step $j-1$: it holds that $S_{j-1} = \mathrm{Cl}(\tilde{\mathcal{A}}_{j-1}(T)) = \mathrm{Cl}(\tilde{\mathcal{A}'}_{j-1})$ is the closure at time-step $j-1$. At time-step j, the original PDT was about to query a linear form ℓ. Suppoe querying ℓ adds the blocks $B_1, B_2, \cdots, B_k$ to ℓ. Then, T' first queries all coordinates in $\mathrm{VARS}(\{B_1, B_2, \cdots, B_k\})$, and then queries $\mathrm{Cl}(\tilde{\mathcal{A}'}_j$. After this time-step j finishes. All necessary invariants are maintained by Lemma 3.26.

To summarize, can describe the behavior of T' as follows:

- At each time-step j, T' queries all coordinates of a set of blocks $T_j \subseteq [n]$. Then, it issues some general linear queries which do not affect the closure.

Now we essentially lift the argument in the proof of Theorem 8.7 in this case. We introduce some notation first:

- Let $S_j = \mathrm{Cl}(\tilde{\mathcal{A}}_{j-1}(T)) = \mathrm{Cl}(\tilde{\mathcal{A}'}_j(x))$ be the closure at time-step j.
- Let $y_j \in \mathbb{F}_2^{\mathrm{VARS}(C_j)}$ be the closure assignment at time-step j. Note that T' knows y_j at the end of time-step j.
- Let $\rho_j = G(y_j)$ and $G_j = G_{\rho_j}$ (the graph obtained by deleting all edges fixed by ρ_j).

We start with the following consequence of Cheeger's inequality.

Lemma B.1. *For all $j \leq n/2000$, the largest connected component of G_j has size at least $n\left(1 - \frac{n}{50d}\right)$.*

Proof. Suppose all connected components of G_j have size at most $n\left(1 - \frac{n}{50d}\right)$. Then, we can pick a subset of the connected components such that their union, S has size in the range $\left[\frac{n}{100d}, n\left(1 - \frac{1}{100d}\right)\right]$. By Cheeger's inequality (Lemma 6.1) the cut $|E(S, V \setminus S)|$ has at least $\frac{n}{500}$ edges. However, the current partial assignment ρ_j can fix at most $|\mathrm{Cl}(A)| + j \leq n/1000$ edges. This is a contradiction. □

Suppose the current partial assignment ρ_j lies in P. By Lemma 8.13, the conditional distribution of $\text{root}(G(x))$, as $x \leftarrow G^{-1}(\mu) \cap \tilde{\mathcal{A}}_j \cap C_{y_j} (= G^{-1}(\mu) \cap \tilde{\mathcal{A}}_j)$ is $o(n^{-20})$ close to uniform on the unique odd component of ρ_j.

We make the PDT T' pay us some amount of coins at each time-step according to the following rules.

- Suppose at time step j, T' queries the sets of coordinates in $B_1, B_2, \cdots, B_k$. Let C be the odd component of ρ_j. In ρ_{j+1}, C splits as

$$C = \tilde{C_1} \cup \tilde{C_2} \cup \cdots \cup \tilde{C_h}$$

  By Lemma B.1, one of the $\tilde{C_j}$s has large size; say $|\tilde{C_1}| \geq n\left(1 - \frac{n}{50d}\right)$. After the responses to the queries are revealed, it is revealed in which component $\text{root}(G(x))$ lies. If it does not lie in $\tilde{C_1}$, the PDT does not pay anything and wins the game. If it lies in $\tilde{C_1}$, the PDT has to pay $|C| - |\tilde{C_1}|$ coins and the game continues.

By Lemma B.1, the PDT never pays more than $\frac{n}{50d}$ coins. So we award the PDT a budget of $\frac{n}{50d}$ coins. We say the PDT *loses the game* when it has to pay some coins but it is broke. Our goal is to show that the PDT loses the game with high probability. We prove the following statement by induction.

Lemma B.2. *Suppose the PDT has c coins remaining and has not won the game yet. Then, the probability it wins the game is at most* $\frac{5c}{n} + \frac{1}{5}$.

Proof. We induct on c. We prove the base case $c = 0$ first. When the PDT has no coins remaining, for the PDT to win the game, whenever it splits the odd component, the root has to lie in one of the smaller components. Suppose the current odd component C splits as $\tilde{C_1} \cup \tilde{C_2} \cup \cdots \cup \tilde{C_h}$ where $|\tilde{C_1}| \geq n\left(1 - \frac{n}{50d}\right)$. By Lemma 8.13, conditioned on the current information obtained (before querying the new blocks), the distribution of $\text{root}(G(x))$ is $o(n^{-20})$ close to uniform on C. So the probability the root does not lie in $|\tilde{C_1}|$ is $\leq \frac{|C| - |\tilde{C_1}|}{|C|} + o(n^{-20}) \leq \frac{1}{5}$. This proves the base case.

Now we prove the induction step. Suppose the PDT has $c > 0$ coins remaining, and at some time-step it splits the current odd component C into $\tilde{C_1} \cup \tilde{C_2} \cup \cdots \cup \tilde{C_h}$ where $|\tilde{C_1}| \geq n\left(1 - \frac{n}{50d}\right)$. By Lemma 8.13, conditioned on current information obtained (before querying the new blocks), the distribution of $\text{root}(G(x))$ is $o(n^{-20})$ close to uniform on C. So the probability the root does not lie in $|\tilde{C_1}|$ is $\leq \frac{|C| - |\tilde{C_1}|}{|C|} + o(n^{-20}) \leq \frac{5(|C| - |\tilde{C_1}|)}{n}$. If the root does not lie in $|\tilde{C_1}|$, the PDT has to pay $|C| - |\tilde{C_1}|$ coins, and by the induction hypothesis it then wins the rest of the game with probability at most $\frac{5(c - (|C| - |\tilde{C}|))}{n} + \frac{1}{5}$. By union bound, the probability the PDT wins the game is at most

$$\leq \frac{5(|C| - |\tilde{C_1}|)}{n} + \frac{5(c - (|C| - |\tilde{C_1}|))}{n} + \frac{1}{5} = \frac{5c}{n} + \frac{1}{5}$$

□

Since the PDT starts with a budget of $\frac{n}{50d}$, the probability it wins the game is at most $\frac{1}{3}$ by Lemma B.2. The desired result now follows.

□

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il