Electronic Colloquium on Computational Complexity, Report No. 18 (2026)

Resolution Width Lifts to Near-Quadratic-Depth Res(⊕) Size

Dmitry Itsykson*[1,2], Vladimir Podolskii†[3], and Alexander Shekhovtsov‡[4]

[1]Ben-Gurion University of the Negev
[2]On leave from Steklov Institute of Mathematics at St. Petersburg
[3]Tufts University
[4]EPFL

February 12, 2026

Abstract

We show that for any unsatisfiable CNF formula φ that requires resolution refutation width at least w, and for any 1-stifling gadget g (for example, $g = \mathrm{MAJ}_3$), (1) every resolution-over-parities (Res(⊕)) refutation of the lifted formula $\varphi \circ g$ of size at most S has depth at least $\Omega(w^2/\log S)$; (2) every Res(⊕) refutation of the lifted formula $\varphi \circ g$ has size $\Omega(w^2)$.

The first result substantially extends and simplifies all previously known lifting theorems for bounded-depth Res(⊕). The lifting result of Itsykson and Knop [22] requires gadgets of logarithmic size and applies only to refutations of depth at most $O(n \log n)$, whereas our result applies to nearly quadratic depth. The liftings of Bhattacharya and Chattopadhyay [9] and of Byramji and Imagliazzo [13] apply to nearly quadratic depth as well, but rely on a much stronger assumption of $(\Omega(n), \Omega(n))$-DT-hardness, which is far less standard than large resolution width.

Our proof combines the random-walk-with-restarts method of Alekseev and Itsykson [4] with a new idea: the random walk is defined relative to the structure of the refutation graph, rather than by a distribution on inputs induced by the formula.

Using this technique, we substantially strengthen the supercritical size-depth tradeoff of Itsykson and Knop [22], both by improving the depth lower bound and by reducing the size of the separating formulas to polynomial in the number of variables, with the latter resolving an open question. In particular, we construct a family of polynomial-size formulas that admit polynomial-size resolution refutations, while any Res(⊕) refutation of depth $o(n^2/\log^4 n)$ necessarily has superpolynomial size.

Our second result yields a pure quadratic lower bound on the size of Res(⊕) refutations, improving upon the previously known near-quadratic lower bound of [13].

1 Introduction

Propositional proof complexity studies propositional proof systems used to certify the unsatisfiability of CNF formulas. One of the central research directions in proof complexity, known as Cook's

*e-mail: dmitrits@gmail.com. Supported by the European Research Council Grant No. 949707.
†e-mail: podolskii.vv@gmail.com
‡e-mail: alex.v.shekhovtsov@gmail.com

ISSN 1433-8092

program, seeks to establish superpolynomial lower bounds on the proof size required by specific proof systems. The ultimate goal of this program is to prove NP $\neq$ coNP, which is equivalent to proving superpolynomial lower bounds for every propositional proof system.

In this paper, we consider two proof systems. The first is *resolution*, the most extensively studied propositional proof system. Resolution proves the unsatisfiability of a formula by deriving new clauses using a single inference rule, which allows one to derive a clause $A \vee B$ from the clauses $A \vee x$ and $B \vee \neg x$. A refutation of a CNF formula φ is a derivation of the empty clause from the clauses of φ. Important complexity measures are the width of a resolution refutation, defined as the maximum number of literals in any clause of the refutation, and the depth, defined as the length of the longest path from an initial clause of the formula to the empty clause.

While many exponential-size lower bounds are known for weak proof systems such as resolution, we lack superpolynomial lower bounds for Frege systems, which include standard propositional proof systems from logic textbooks. A Frege derivation is a sequence of Boolean formulas; each of them is either an axiom or is obtained from the previous by a set of sound and implicationally complete inference rules. Proving Frege lower bounds is often compared to proving Boolean formula/circuit lower bounds for explicit Boolean functions, and both seem intractable. However, progress has been made in restricted settings. An exponential lower bound for constant-depth circuits computing parity was proven in the 1980s [18, 1]. Ajtai later used a similar approach to prove a superpolynomial lower bound for bounded-depth Frege systems [2]. Razborov and Smolenski proved lower bounds for constant-depth circuits with $\neg$, $\vee$, $\wedge$, and MOD_p gates in 1987 [30, 29]. The analogous problem of proving a lower bound for constant-depth Frege systems using $\neg$, $\vee$, $\wedge$ and MOD_p gates (denoted $\mathrm{AC}^0[p]$-Frege) is open for all $p > 1$.

The second proof system considered in this paper is *resolution over parities* (Res($\oplus$)), an extension of resolution that permits reasoning modulo 2. In contrast to standard resolution, Res($\oplus$) operates with linear clauses, which are disjunctions of linear equations over $\mathbb{F}_2$. Its resolution rule allows one to infer $C \vee D$ from the premises $C \vee (f = 0)$ and $D \vee (f = 1)$, where f is a linear form. In addition, Res($\oplus$) includes a weakening rule, which allows one to derive a linear clause D from a clause C whenever C semantically implies D. The depth of a Res($\oplus$) refutation is the maximum number of resolution steps along a path from an initial clause to the empty clause.

Res($\oplus$) is the weakest known subsystem of $\mathrm{AC}^0[2]$-Frege for which superpolynomial lower bounds are still open. This makes the study of lower bounds for Res($\oplus$) a particularly important direction.

1.1 Superpolynomial Lower Bounds for Subsystems of Res($\oplus$)

The first lower bounds for tree-like Res($\oplus$) were proved by Itsykson and Sokolov [24, 25]. Since then, a number of works have established exponential lower bounds for tree-like Res($\oplus$) on classical combinatorial formulas [20, 21, 23, 27, 13].

Chattopadhyay, Mande, Sanyal, and Sherif [14], and independently Beame and Kroth [6], introduced a lifting approach for proving lower bounds in tree-like Res($\oplus$).

Given a CNF formula $\varphi(y_1, y_2, \ldots, y_n)$ and a Boolean function $g : \{0,1\}^\ell \to \{0,1\}$, called a *gadget*, the *lifted formula* $\varphi \circ g$ is defined as the CNF encoding of

$$\varphi\big(g(x_{1,1}, \ldots, x_{1,\ell}), \ldots, g(x_{n,1}, \ldots, x_{n,\ell})\big),$$

where each variable y_i in φ is replaced by $g(x_{i,1}, \ldots, x_{i,\ell})$ over fresh variables $x_{i,1}, \ldots, x_{i,\ell}$.

Chattopadhyay, Mande, Sanyal, and Sherif [14] also introduced the notion of *k-stifling gadgets*. A Boolean function $g : \{0,1\}^\ell \to \{0,1\}$ is called a k-stifling gadget if, for every $a \in \{0,1\}$ and every

choice of $\ell - k$ input variables, there exists an assignment to these $\ell - k$ variables such that the value of g is fixed to a, regardless of the values of the remaining k variables.

They [14] showed that if every resolution refutation of a formula φ has depth at least h, and g is a k-stifling gadget, then any tree-like Res($\oplus$) refutation of the lifted formula $\varphi \circ g$ must have size at least 2^{kh}.

Efremenko, Garlik, and Itsykson [15] made the first progress beyond tree-like Res($\oplus$) by proving an exponential lower bound for (bottom-)regular Res($\oplus$). Building on this result, Alekseev and Itsykson [4] established an exponential lower bound for a stronger model—namely, Res($\oplus$) refutations of depth at most $n \log \log n$, where n is the number of variables. Their proof introduced a technique based on random walks with restarts, which has since become standard in this area. Subsequently, Efremenko and Itsykson [16] presented an alternative analysis of the same approach, improving the depth lower bound to $n \log n$. In all these results, the hard instance is a Tseitin formula over an $O(\log n)$-degree expander, lifted by an arbitrary 2-stifling gadget.

Itsykson and Knop [22] proved the following lifting theorem.

Theorem 1.1 ([22])**.** Suppose that every resolution refutation of φ has either width at least w or depth at least h. Let s be an integer such that $h \geq s^2 w$. Then any Res($\oplus$) refutation of $\varphi \circ \oplus_s \circ \mathrm{MAJ}_5$ has either size at least 2^w or depth $\Omega(s^2 w)$.

Theorem 1.1 implies that any CNF formula with resolution width $\Omega(n)$ can be lifted by an $O(\log n)$-size gadget to a formula for which every Res($\oplus$) refutation has either exponential size or depth $\Omega(n \log n)$.

Another consequence of Theorem 1.1 is that, when combined with the supercritical width–depth tradeoff for resolution due to Buss and Thapen [10], it yields a *supercritical tradeoff* between depth and Res($\oplus$) size. Specifically, there exists a family of formulas ψ_n on n variables, of size $n^{O(\log^2 n)}$, that admit resolution refutations of size $n^{O(\log^2 n)}$, yet for which any Res($\oplus$) refutation of size at most $2^{n/\log^4 n}$ must have depth $\Omega(n \log n)$. This depth bound exceeds the trivial worst-case upper bound of n, which is why such a tradeoff is termed *supercritical*. Moreover, applying Theorem 1.1 with larger values of s shows that Depth-$n^{4/3-\varepsilon}$ Res($\oplus$) does not polynomially simulate resolution for any $\varepsilon > 0$.

All of the aforementioned papers [4, 16, 22] apply the random-walk-with-restarts method under the uniform distribution on inputs. This choice leads to an exponentially small success probability, which in turn forces the random walk to terminate after only a small number of rounds, thereby limiting the achievable depth lower bounds.

Bhattacharya and Chattopadhyay [9] were the first to apply the random-walk-with-restarts technique under a non-uniform input distribution. This innovation yields a 2^{n^ε} lower bound for the size of Res($\oplus$) refutations of nearly quadratic depth $n^{2-\varepsilon}$. Their hard instance is obtained by lifting a formula with the IP gadget of size $O(\log n)$. The base formula is required to satisfy an $(\Omega(n), \Omega(n))$-DT hardness property; although this property is somewhat technical, it is known to hold for Tseitin formulas over constant-degree expanders.

In recent work, Byramji and Impagliazzo [13], building on ideas from lifting theorems for parity decision trees [14, 6, 28, 12], obtained several additional results. Among them a 2^{n^ε} lower bound on the size of Res($\oplus$) refutations of the binary pigeonhole principle BPHP_n^{n+1} of depth $n^{2-\varepsilon}$. Their main lifting contribution is an improvement in the gadget size: it suffices to lift an $(\Omega(n), \Omega(n))$-DT-hard formula using an arbitrary 1-stifling gadget. When applied to Tseitin formulas, this yields lower

bounds for formulas of size $O(n)$, meaning that the refutation depth is almost quadratic in the formula size, rather than in the number of variables as in previous works.

In a recent paper, Alekseev and Gaevoi [3] introduced a new family of formulas, the *Constrained Bit Pigeonhole Principle*, and proved a lower bound of $2^{n^{\Omega(\varepsilon)}}$ for Res($\oplus$) refutations of depth at most $n^{2-\varepsilon}$. Using these formulas, the authors were able to apply the random-walk-with-restarts framework with the uniform distribution over inputs.

Recently, Efremenko and Itsykson [17] established that for every $\varepsilon > 0$, a lower bound of $2^{n(1-\varepsilon)}$ holds for Depth-n Res($\oplus$) refutations, where n denotes the number of variables in the refuted formula. Consequently, this lower bound is essentially as strong as what is predicted by the Strong Exponential Time Hypothesis (SETH). Their proof relies on a refinement of the random-walk-with-restarts method, combined with a lifting argument from formulas of extremely large resolution width.

1.2 Polynomial Lower Bounds for Unrestricted Res($\oplus$)

Khaniki [26] proved an almost quadratic lower bound $\Omega\left(\frac{n^2}{\log^{\log\log n} n}\right)$ for a dag-like version of Res($\oplus$) with *syntactic weakening rules*. It remains unclear whether this lower bound persists for the semantic weakening rules that have since become standard. Although these variants of Res($\oplus$) are polynomially related [25], this relationship does not yield polynomial lower bounds.

Prior to our work, the only known superlinear lower bound for the standard version of Res($\oplus$) (i.e., with semantic weakening) followed from the results of Byramji and Impagliazzo [13]. Specifically, their size–depth tradeoff implies an $\Omega\left(\frac{n^2}{\log n}\right)$ lower bound on the size of Res($\oplus$) refutations of Tseitin formulas over constant-degree expanders lifted by the MAJ_3 gadget.

1.3 Our Goals

The main goal of this paper is to identify a standard complexity measure of CNF formulas that enables lifting to size lower bounds for Res($\oplus$) refutations of depth up to $n^{2-\varepsilon}$. Existing lifting results by Bhattacharya and Chattopadhyay [9] and by Byramji and Impagliazzo [13] rely on a rather strong and nonstandard assumption of $(\Omega(n), \Omega(n))$-DT hardness. Our objective is to replace this assumption with a more natural and well-studied measure.

Motivated by the lifting result of Itsykson and Knop [22], which applies to smaller depths, and by the result of Alekseev and Itsykson [4], which lifts resolution width to width in Res($\oplus$), we aim to show that resolution width—defined as the minimum possible width of a resolution refutation—can serve as an appropriate source of hardness.

Resolution width is a standard and extensively studied measure in proof complexity. Since the seminal work of Ben-Sasson and Wigderson [8], which showed that resolution width lower bounds of $n^{1/2+\varepsilon}$ imply exponential lower bounds on resolution size, proving width lower bounds has become the standard approach for establishing size lower bounds in resolution. Atserias and Dalmau [5] provided a game characterization of resolution width that offers a general framework for proving such lower bounds and, moreover, implies that clause space in resolution is also lower bounded by resolution width. This characterization readily implies that any $(\Omega(n), \Omega(n))$-DT hard formula has resolution width at least $\Omega(n)$.

Beck and Impagliazzo [7] demonstrated that resolution width can be lifted via the parity gadget $\oplus_2$ to resolution size: if a formula φ requires resolution width w, then the composed formula $\varphi \circ \oplus_2$

requires resolution size $2^{\Omega(w)}$.

Finally, Garg, Göös, Kamath, and Sokolov [19] showed that resolution width lifts via an indexing gadget to size lower bounds in cutting planes and in DAG-like communication protocols. The latter result, in particular, implies lower bounds for a wide range of proof systems as well as for the size of monotone Boolean circuits.

Another goal of the paper is to address the near-quadratic lower bound for Res($\oplus$) established via the size–depth tradeoff of [13]. We aim to obtain a direct proof of such a lower bound, or possibly of a stronger one.

1.4 Our Contributions

We prove the following lifting theorem.

Theorem 1.2 (Theorem 6.1)**.** Assume that every resolution refutation of an unsatisfiable CNF formula φ has either width at least w, or depth at least h. Let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-stifling gadget, where ℓ is a constant. Then any Res($\oplus$) refutation of $\varphi \circ g$ of size at most S has depth at least $\Omega\left(\min\{h, w^2/\log S\}\right)$.

A special case of Theorem 1.2 yields a lifting theorem from resolution width.

Corollary 1.3. Assume that every resolution refutation of an unsatisfiable CNF formula φ has width at least w. Let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-stifling gadget, where ℓ is a constant. Then any Res($\oplus$) refutation of $\varphi \circ g$ of size at most S has depth at least $\Omega\left(w^2/\log S\right)$.

Corollary 1.3 implies that, starting from formulas requiring resolution width $\Omega(n)$, one obtains 2^{n^ε} lower bounds for Depth-$n^{2-\varepsilon}$ Res($\oplus$) refutations, as in [9, 13]. Moreover, if the starting formula is an $O(1)$-CNF and has $O(n)$ clauses, then, as in [13], one can obtain formulas for which the size–depth tradeoff yields nearly quadratic depth in terms of the *formula size*, rather than the number of variables. The advantage of our result compared to [9, 13] is that it relies on a standard complexity measure and applies to a wide range of formulas, thanks to the numerous known lower bounds on resolution width.

Let us compare Theorem 1.2 with the main result of [22], namely Theorem 1.1. First, Theorem 1.2 relies on a significantly simpler gadget: any 1-stifling function, as opposed to a composition of parity with a 2-stifling gadget (namely MAJ_5). Second, Theorem 1.2 yields larger depth lower bounds without a substantial increase in the size of the formula.

Combining Theorem 1.2 with the supercritical width–depth tradeoff for resolution due to Buss and Thapen [11], we obtain the following theorem.

Theorem 1.4 (Theorem 6.4)**.** There is a family of formulas ψ_n from n variables of size $\mathrm{poly}(n)$ such that ψ_n has polynomial size resolution refutation for any $S > 0$, any Res($\oplus$) refutation of size S has depth at least $\Omega(n^2/(\log^2 n \log S))$.

Theorem 1.4 improves the result of [22] in two respects. First, the formulas ψ_n have polynomial size and admit polynomial-size resolution refutations, whereas in [22] the corresponding formulas were only of quasipolynomial size. This resolves an open question posed in [22]. Second, the theorem applies to a wider range of depths, which can be used to obtain stronger separation between resolution and bounded depth Res($\oplus$).

Corollary 1.5 (Corollary 6.5)**.** If $d(n) = o(n^2/\log^4 n)$, then Depth-$d(n)$ Res($\oplus$) does not polynomially simulates resolution.

Size lower bound. Applying Corollary 1.3 to an $O(1)$-CNF formula φ with n variables and $O(n)$ clauses that requires resolution width $\Omega(n)$, and choosing $S = n^2$ and $g = \mathrm{MAJ}_3$, we obtain that every Res($\oplus$) refutation of $\varphi \circ \mathrm{MAJ}_3$ (which has size $O(n)$) with size at most n^2 must have depth at least $\Omega\left(\frac{n^2}{\log n}\right)$. The same lower bound can also be derived from the results of [13]. We strengthen this conclusion by obtaining a *pure* quadratic lower bound. This improvement holds for the same class of formulas as above and is established by the following theorem.

Theorem 1.6 (Theorem 7.4)**.** Let φ be an unsatisfiable CNF formula such that every resolution refutation of φ has width at least w. Let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-stifling gadget. Then any Res($\oplus$) refutation of $\varphi \circ g$ has size at least $\frac{w(w+1)}{2}$.

1.5 Our Techniques

We always use the top-down definition of Res($\oplus$) as a parity decision DAG [15]. A *parity decision DAG* refuting an unsatisfiable CNF formula φ is a directed acyclic graph with a single source and several sinks, satisfying the following properties:

- Each node v of the DAG is labeled with an $\mathbb{F}_2$-linear system Φ_v over the variables of φ.
- The source is labeled with the empty system (i.e., identically true).
- For every sink v, there exists a clause C of φ such that Φ_v is inconsistent with C.
- Every non-sink node v is additionally labeled with a linear form f_v and has two children v_0 and v_1. The edge (v, v_0) is labeled by the equation $f_v = 0$, and the edge (v, v_1) is labeled by $f_v = 1$. Moreover, for each $\alpha \in \{0,1\}$, the system $\Phi_v \wedge (f_v = \alpha)$ semantically implies Φ_{v_α}.

Quadratic size lower bound. Let us start with presenting the proof idea of Theorem 1.6.

Let $\varphi(y_1, \ldots, y_m)$ be an unsatisfiable CNF formula whose resolution width is at least w, and let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-stifling gadget. Recall that the lifted formula $\varphi \circ g$ is over the variable set $\{x_{i,j} \mid i \in [m],\, j \in [\ell]\}$. Suppose that G is a parity decision DAG refuting $\varphi \circ g$. Our goal is to show that the size of G is at least $\frac{w(w+1)}{2}$.

We prove this by induction on w. For the induction step, it suffices to construct a restriction ρ such that:

- the restricted DAG $G|_\rho$ refutes $\varphi' \circ g$, where φ' is an unsatisfiable formula of resolution width at least $w - 1$; and
- the restriction ρ falsifies at least w linear systems labeling nodes of G. Consequently, the size of $G|_\rho$ is smaller than the size of G by at least w.

Once such a restriction is obtained, the induction follows immediately.

Fix an index $i \in [m]$ and choose $a \in \{0,1\}$ such that the restricted formula $\varphi|_{y_i=a}$ has resolution width at least $w - 1$. Since g is 1-stifling, there exists a partial assignment σ that assigns Boolean values to all variables $\{x_{i,j} \mid j \in [\ell]\}$ except for a single variable x_{i,j_0}, in such a way that the value of $g(x_{i,1}, \ldots, x_{i,\ell})$ is fixed to a, independently of the value of x_{i,j_0}.

Observe that $G|_\sigma$ is a parity decision DAG refuting $\varphi|_{y_i=a} \circ g$. Hence, we may assume that the variable x_{i,j_0} does not appear in the clauses falsifying at the sinks of $G|_\sigma$. Therefore, we may substitute an arbitrary affine form h for x_{i,j_0}—that is, define

$$\rho := \sigma \cup \{x_{i,j_0} := h\},$$

and the resulting DAG $G|_\rho$ still refutes $\varphi|_{y_i=a} \circ g$.

Hence, it suffices to find an affine equation that is semantically implied by at least w linear systems labeling nodes of G. Once such an equation is identified, we choose the variable x_{i,j_0} appearing in it and substitute for it an affine form that falsifies the equation, thereby eliminating all these nodes.

By the result of Alekseev and Itsykson [4], any parity decision DAG refuting $\varphi \circ g$ contains a node v whose associated linear system Φ_v has rank at least w. Using standard properties of parity decision DAGs, one can show that along the path from the source to v there exists an affine equation that is implied by at least w of the linear systems labeling nodes on this path. This completes the induction step.

Lower bound for bounded-depth Res($\oplus$)**.** Next, we describe our proof strategy for our main lifting result. For simplicity, we concentrate on the proof of Corollary 1.3.

Again, consider a parity decision DAG G of size S refuting $\varphi \circ g$.

The proof goes by induction on the resolution width w of the CNF φ. The key step of the proof is to find a node v in G on depth at least $w/6$ and to substitute some of the variables of $\varphi \circ g$ by affine functions of the remaining variables in such a way that the subgraph of G with the root in v refutes $\varphi' \circ g$ and the resolution width of φ' is at least $w - O(\log S)$. Once we show how to find such a v, the result follows easily.

To find the vertex v, similarly to the previous papers, we adopt the random-walk-with-restart technique. We follow a random path in G starting from the root for $O(w)$ steps. We show that if the rank of the linear system Φ_v in the final vertex of the path is d, then this vertex can be reached by the random path only with probability $2^{-\Omega(d)}$ (here, for simplicity we assume that the size ℓ of the gadget g is constant). Thus, if all final vertices have linear systems of rank larger than $O(\log S)$, then there are more than S of them, which contradicts the restriction on the size of G. As a result, there is a vertex v that has a system of rank at most $O(\log S)$. We show that in this case, we can fix $O(\log S)$ input variables of φ to obtain a refutation of $\varphi' \circ g$ with the desired bound on the resolution width of φ'.

To achieve that, as in the previous case (that is, the quadratic size lower bound), we fix in each block of variables of $\varphi \circ g$ all but one variable to 0/1 constants. These values are chosen—again as in the previous case—so as to reduce the width of φ by at most one per block. The remaining unfixed variables are then substituted by affine functions chosen to satisfy the linear system Φ_v. A sufficient condition ensuring that such a restriction satisfies the linear system is formulated in terms of *closure* (see Lemma 4.5 for details). By the properties of closure, it suffices to verify this condition for the linear system written along the edges of the path. This is the standard use of closure, dating back to its introduction in [15].

The key novelty of our argument is the way we pick a random path. Previous papers picked a random input and followed the path corresponding to this input. Instead, we gradually fix the input during the random walk: on each step of the walk, we fix some bits in order to be able to move to the next vertex. For this, in each intermediate vertex v of the path, we consider the linear form f_v queried in v. We pick some i such that $\{x_{i,j} \mid j \in [\ell]\}$ contains a variable x_{ij} of f. We consider the variable y_i of φ and pick the value $a \in \{0,1\}$ for it such that, informally speaking, fixing $y_i = a$ decreases the resolution width of φ by at most 1. Then we flip a coin and do one of the following with probabilities 1/2:

(1) Assign the values to variables in $\{x_{i,j} \mid j \in [\ell]\}$ uniformly at random among $g^{-1}(a)$.

(2) Assign all variables of $\{x_{i,j} \mid j \in [\ell]\}$ except x_{ij} in such a way that g evaluates to a regardless of the value of x_{ij} (this is possible since g is 1-stifling). Then we fix x_{ij} to be the linear combination of other variables of f in such a way that f evaluates to 0 and 1 with probabilities 1/2.

Next, we repeat the procedure, in the second case in the next vertex of G, and in the first case potentially in the same vertex of G.

The combination of these two actions allows us to show that the probability that the path reaches any given vertex labeled by a linear system with large rank is small. Here we use properties of 1-stifling gadget again (see Lemma 2.5). Additionally, the second action ensures that the random path makes many steps in G.

Organization. The rest of the paper is organized as follows. In Section 2, we give the necessary definitions and prove the basic properties of 1-stifled gadgets. In Section 3, we study proof-complexity properties of CNF formulas via families of games and winning strategies, which are typically represented as sets of partial assignments. In Section 4, we analyze the application of affine restrictions to parity decision DAGs computing lifted formulas, and we present a sufficient condition under which a linear system can be satisfied by an affine restriction of small size. In Section 5, we discuss the construction of the random paths and their properties. In Section 6, we combine everything together to obtain our lifting result and its corollaries. Finally, in Section 7, we prove a quadratic size lower bound.

2 Preliminaries

Throughout this paper, all scalars are from the field $\mathbb{F}_2$. Let X be a set of variables taking values in $\mathbb{F}_2$. A linear form in variables from X is a homogeneous linear polynomial over $\mathbb{F}_2$ in variables from X or, in other words, a polynomial $\sum_i^n x_i a_i$, where $x_i \in X$ is a variable and $a_i \in \mathbb{F}_2$ for all $i \in [n]$. An affine form is an arbitrary linear polynomial over $\mathbb{F}_2$, i.e. $\sum_i^n x_i a_i + a_0$, where $x_i \in X$ is a variable and $a_i \in \mathbb{F}_2$ for all $i \in \{0, 1, \ldots, n\}$. A linear equation is an equality $f = a$, where f is a linear form and $a \in \mathbb{F}_2$.

2.1 CNF Formulas

A *literal* is either a propositional variable or its negation. For a propositional variable x, we write $x^0 := \neg x$ and $x^1 := x$.

A *clause* is a disjunction of literals. The empty clause is identified with the constant false. A *CNF formula* is a conjunction of clauses, and the empty conjunction is identified with the constant true.

Observe that a conjunction of CNF formulas is again a CNF formula. Let $\phi_1, \phi_2, \ldots, \phi_k$ be CNF formulas, and assume that for each $i \in [k]$,

$$\phi_i = \bigwedge_{j \in K_i} C_{i,j},$$

where each $C_{i,j}$ is a clause. We define the *standard CNF representation* of the disjunction $\bigvee_{i=1}^k \phi_i$

to be the CNF formula

$$\bigwedge_{(j_1,\ldots,j_k)\in K_1\times\cdots\times K_k} (C_{1,j_1} \vee C_{2,j_2} \vee \cdots \vee C_{k,j_k})\,. \tag{1}$$

2.2 Partial Assignments and Affine Restrictions

Affine restrictions. By an *affine restriction* we mean a set of assignments of the form

$$\rho = \{x_1 := f_1, x_2 := f_2, \ldots, x_k := f_k\},$$

where $x_1, \ldots, x_k$ are distinct variables, each $f_1, \ldots, f_k$ is an $\mathbb{F}_2$-affine form, and none of the variables $x_1, \ldots, x_k$ occurs in any of the forms $f_1, \ldots, f_k$. The set $\{x_1, \ldots, x_k\}$ is called the *support* of ρ.

Partial assignments. An affine restriction ρ is called *plain*, or a *partial assignment*, if all affine forms f_i from the right hand sides are 0/1-constants.

We say that a partial assignment ρ *satisfies* a literal x^a if $x := a \in \rho$, and *falsifies* x^a if $x := 1 - a \in \rho$.

Let C be a clause and let ρ be a partial assignment. The restricted clause $C|_\rho$ is defined as follows. If ρ satisfies at least one literal of C, then $C|_\rho$ is the constant 1, and we say that ρ satisfies C. Otherwise, $C|_\rho$ is obtained from C by deleting all literals falsified by ρ. We say that ρ *falsifies* C if $C|_\rho$ is the empty clause.

For a CNF formula ϕ and a partial assignment ρ, we define the restricted CNF formula $\phi|_\rho$ as follows. The formula $\phi|_\rho$ is the constant 0 if ρ falsifies any clause of ϕ; otherwise,

$$\phi|_\rho \;=\; \bigwedge_{C\in\phi\,:\,\rho \text{ does not satisfy } C} C|_\rho.$$

2.3 Resolution

Let φ be an unsatisfiable CNF formula. A resolution refutation of φ is a sequence of clauses $C_1, C_2, \ldots, C_s$ such that C_s is the empty clause (i.e., identically false) and for every $i \in [s]$ the clause C_i is either a clause of φ or is obtained from previous clauses by the *resolution rule* that allows us to derive a clause $C \vee D$ from clauses $C \vee x$ and $D \vee \neg x$.

The *size* of a resolution refutation is the number of clauses in it. The *depth* of a resolution refutation is the length of the longest path between the empty clause and the clause of the original formula. The *width* of a resolution refutation is the maximal size of a clause from the refutation, where the size of a clause is the number of literals it contains. The resolution width of an unsatisfiable CNF formula φ is the minimal possible width over all resolution refutations of φ.

2.4 Resolution Over Parities

A *linear clause* is a disjunction of $\mathbb{F}_2$-linear equations: $\bigvee_{i=1}^{t}(f_i = a_i)$. Note that over $\mathbb{F}_2$ a linear clause $\bigvee_{i=1}^{t}(f_i = a_i)$ may be represented as the negation of a linear system: $\neg \bigwedge_{i=1}^{t}(f_i = a_i + 1)$.

Now we define the proof system resolution over parities (Res($\oplus$)) [25].

Let φ be an unsatisfiable CNF formula. A Res($\oplus$) refutation of φ is a sequence of linear clauses C_1, C_2, …, C_s such that C_s is the empty clause (i.e., identically false) and for every $i \in [s]$ the clause C_i is either a clause of φ or is obtained from previous clauses by one of the following inference rules:

– *Resolution rule* allows us to derive a linear clause $C \vee D$ from linear clauses $C \vee (f = a)$ and $D \vee (f = a + 1)$.
– *Weakening rule* allows us to derive from a linear clause C any linear clause D in the variables of φ that semantically follows from C (i.e., any assignment satisfying C also satisfies D).

The *size* of a Res($\oplus$) refutation is the number of linear clauses in it. The *depth* of a Res($\oplus$) refutation is the maximal number of resolution rules applied on a path between a clause of the initial formula and the empty clause. Note that weakening rules are not counted in the definition of the depth. The *width* of a Res($\oplus$) refutation is the maximal rank of the negation of a linear clause from the refutation.

Remark 2.1. A resolution refutation of a formula φ is a special case of a Res($\oplus$) refutation, where all linear clauses are plain (i.e., disjunctions of literals).

For any function $f(n)$, we denote by Depth-$f(n)$ Res($\oplus$) the subsystem of Res($\oplus$) consisting of refutations with depth at most $f(n)$, where n is the number of variables in the formula being refuted.

For a linear clause C we denote by $L(C)$ the set of linear forms that appear in C; i.e. $L\left(\bigvee_{i=1}^{t}(f_i = a_i)\right) = \{f_1, f_2, \ldots, f_t\}$. The same notation we use for linear systems: if Ψ is a $\mathbb{F}_2$-linear system, $L(\Psi)$ denotes the set of all linear forms from Ψ.

2.5 Parity Decision DAG

We say that an $\mathbb{F}_2$-linear system Φ_1 *semantically implies* an $\mathbb{F}_2$-linear system Φ_2 if every assignment satisfying Φ_1 also satisfies Φ_2. We denote semantic implication by $\Phi_1 \models \Phi_2$.

A *parity decision DAG* (also known as an affine DAG) is a directed acyclic graph with a single source and several sinks, satisfying the following properties:

– Each node v of the DAG is labeled with an $\mathbb{F}_2$-linear system Φ_v.
– Every non-sink node v is additionally labeled with a linear form f_v and has two children v_0 and v_1. The edge (v, v_0) is labeled by the equation $f_v = 0$, and the edge (v, v_1) is labeled by $f_v = 1$. Moreover, for each $\alpha \in \{0, 1\}$, the system $\Phi_v \wedge (f_v = \alpha)$ semantically implies Φ_{v_α}.

We say that a parity decision DAG refutes an unsatisfiable CNF formula φ if all linear systems and linear forms used as labels are over Vars(φ) and

– The source is labeled with the empty system (i.e., identically true).
– For every sink v, there exists a clause C of φ such that Φ_v is inconsistent with C.

The size of a parity decision DAG is the number of nodes, the depth is the length of the longest source-to-sink path, and the width is the maximal rank of the linear systems labeling nodes of the DAG.

Similarly to the case of resolution, it is known [15] that every Res($\oplus$) refutation of φ can be efficiently transformed into a parity decision DAG refuting φ without increasing its size, width, and depth.

Lemma 2.2 ([15])**.** Let G be a parity decision DAG, and let u and v be two nodes such that there exists a path p from u to v. Let Ψ_p denote the linear system consisting of the conjunction of all equations labeling the edges along p. Then $\Phi_u \wedge \Psi_p \models \Phi_v$.

2.6 Lifting Settings

We assume that $Y = \{y_1, y_2, \ldots, y_m\}$ is the set of variables of an unlifted formula φ.

Let $g : \{0,1\}^\ell \to \{0,1\}$ be a Boolean function that we refer to as a gadget. We now define a lifted version of the formula φ by a gadget g. The set of variables X of the lifted formula is partitioned into m blocks, where each block corresponds to a variable from Y. The variables in X are indexed by two indices, where the first index indicates the block: $X = \{x_{i,j} \mid i \in [m]\ j \in [\ell]\}$.

We define the lifted formula $\varphi \circ g$ as follows. For every $i \in [m]$, we substitute the variable y_i by $g(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell})$. We then convert the resulting formula into CNF in a clause-wise manner. That is, for each clause C of φ, we perform the substitution to obtain a formula $C \circ g$, convert it to CNF, and finally take the conjunction of all resulting clauses over all C.

We now explain how to convert $C \circ g$ into CNF. Let $C = \bigvee_{s=1}^{t} y_{i_s}^{a_s}$ be a clause. Then $C \circ g = \bigvee_{s=1}^{t} g^{a_s}(x_{i_s,1}, x_{i_s,2}, \ldots, x_{i_s,\ell})$. Assume that for every $a \in \{0,1\}$, we fix a CNF representation of the Boolean function g^a. Under this assumption, $C \circ g$ is represented in CNF using the standard encoding of a disjunction of CNF formulas, as specified in (1).

An affine restriction ρ over lifted variables is called *block-respecting* if, for every block $i \in [m]$, ρ is either undefined on all variables of block i or defined on all variables of block i.

The *block-size* of ρ is the number of blocks on which ρ is defined; we denote it by $\mathrm{bsize}(\rho)$.

Let $g : \{0,1\}^\ell \to \{0,1\}$ be a gadget. For every block-respecting partial assignment τ to the lifted variables X, we define the *induced assignment* $\mathrm{induced}_g(\tau)$ as a partial assignment σ to the unlifted variables in the same set of blocks where τ is defined, such that for each $i \in [m]$ and $j \in [\ell]$, $\sigma(y_{i,j}) = g\big(\tau(x_{i,1}), \tau(x_{i,2}) \ldots \tau(x_{i,\ell}))\big)$.

2.7 Stifling and Subspace-Resilient Gadgets

In this section, we introduce and discuss the properties of gadgets that are needed for the lifting construction.

Definition 2.3 (1-stifling function [14])**.** A function $g : \{0,1\}^\ell \to \{0,1\}$ is called 1-*stifling* if for every index $i \in [\ell]$ and every output value $a \in \{0,1\}$, there exists an input $x \in \{0,1\}^\ell$ such that for all inputs $y \in \{0,1\}^\ell$ that agree with x on all coordinates in $[\ell] \setminus \{i\}$, we have $g(y) = a$.

Our main example of 1-stifling gadget is the 3-bit majority function $\mathrm{MAJ}_3 : \{0,1\}^3 \to \{0,1\}$. It is straightforward to verify that this function is 1-stifling.

We also require a property ensuring that the preimages of g are not fully correlated with any affine subspace of codimension 1.

Definition 2.4 (1-subspace-resilient function)**.** A function $g : \{0,1\}^\ell \to \{0,1\}$ is called 1-*subspace-resilient* if for every affine subspace $W \subseteq \{0,1\}^\ell$ of codimension 1 and every $y \in \{0,1\}$, the preimage $g^{-1}(y)$ is not contained in W.

Next we observe a simple quantitative version of subspace resilience.

Lemma 2.5. Let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-subspace-resilient function, $y \in \{0,1\}$ and $W \subseteq \{0,1\}^\ell$ be an affine subspace of codimension 1. Pick element x of $g^{-1}(y)$ uniformly at random. Then

$$\Pr[x \notin W] \geq 2^{-l}.$$

Proof. Since g is 1-subspace-resilient, we have $g^{-1}(y) \not\subseteq W$, and hence $|g^{-1}(y) \cap W| \leq |g^{-1}(y)| - 1$. Therefore,

$$\Pr[x \in W] = \frac{|g^{-1}(y) \cap W|}{|g^{-1}(y)|} \leq 1 - \frac{1}{|g^{-1}(y)|} \leq 1 - 2^{-\ell},$$

where the last inequality follows from the trivial bound $|g^{-1}(y)| \leq 2^{\ell}$. □

The following lemma establishes that 1-stifling functions are necessarily 1-subspace-resilient:

Lemma 2.6. Let $g : \{0,1\}^{\ell} \to \{0,1\}$ be a 1-stifling function. Then g is 1-subspace-resilient.

Proof. Any affine subspace $W \subseteq \{0,1\}^{\ell}$ of codimension 1 can be described by a linear equation

$$a_1 x_1 + \cdots + a_\ell x_\ell = b \pmod 2,$$

where not all coefficients a_i are zero. Suppose, towards a contradiction, that there exists $y \in \{0,1\}$ such that the entire preimage $g^{-1}(y)$ is contained in W. Let $i \in [\ell]$ be an index with $a_i = 1$. By 1-stiflingness, there exists an input $x \in \{0,1\}^{\ell}$ such that $g(z) = y$ for all inputs z agreeing with x on all coordinates except possibly i. In particular, both x and $x^{\oplus i}$ belong to $g^{-1}(y)$, where $x^{\oplus i}$ denotes x with the i-th bit flipped. However, exactly one of x and $x^{\oplus i}$ satisfies the defining equation of W, since flipping the i-th bit toggles the left-hand side. This contradicts the assumption that $g^{-1}(y) \subseteq W$. Hence, g is 1-subspace-resilient. □

Corollary 2.7. MAJ_3 is 1-subspace-resilient function.

2.8 Safe Sets and Closure

In our proofs, we use the notion of *closure*, introduced in [15], which we review and formalize in this section.

Let $X = \{x_{i,j} \mid i \in [m],\ j \in [\ell]\}$ be the set of lifted variables.

Let F be a finite set of $\mathbb{F}_2$-linear forms over a set of variables X. By $\langle F \rangle$ we denote the linear span of the set F. We say that F is *safe* if there exists a basis $f_1, f_2, \ldots, f_k$ of $\langle F \rangle$ and variables $x_{i_1,j_1},\ x_{i_2,j_2},\ \ldots,\ x_{i_k,j_k}$ such that $i_1, i_2, \ldots, i_k$ are pairwise distinct elements of $[m]$ and, for every $t \in [k]$, the variable x_{i_t,j_t} appears in f_t with a nonzero coefficient and does not appear in f_s for $s \neq t$.

For a set of blocks $S \subseteq [m]$ and a linear form f, let $f[\backslash S]$ denote the linear form obtained from f by setting to zero (equivalently, removing) all variables whose blocks belong to S. For a set of linear forms F, define

$$F[\backslash S] = \{f[\backslash S] \mid f \in F\}.$$

A *closure* of a set of linear forms F is an inclusion-minimal set $S \subseteq [m]$ such that $F[\backslash S]$ is safe.

Lemma 2.8 ([15])**.** For any F, its closure is unique.

Since the closure of F is unique, we denote it by $\mathrm{Cl}(F)$.

We also need the following properties of closure.

Lemma 2.9 ([15])**.** Closure satisfies the following properties:

(1) (Monotonicity) If $F \subseteq G$, then $\mathrm{Cl}(F) \subseteq \mathrm{Cl}(G)$;
(2) (Span invariant) If $\langle F \rangle = \langle G \rangle$, then $\mathrm{Cl}(F) = \mathrm{Cl}(G)$;
(3) (Size bound) $|\mathrm{Cl}(F)| + \dim\langle F[\setminus \mathrm{Cl}(F)]\rangle \leq \dim\langle F \rangle$.

3 Sets of Good Assignments

In this section, we define certain sets of partial assignments for unlifted formulas that capture their essential properties. We first give a general, abstract definition; in Subsection 3.1, we give a specialization that characterizes the width-depth tradeoff in resolution.

Let φ be an unsatisfiable CNF formula over the set of variables $Y = \{y_i \mid i \in [m]\}$.

Definition 3.1 (Set of w-good assignments)**.** For a natural number w, a set $\mathcal{H}$ is called a *set of w-good assignments* for φ if $\mathcal{H}$ consists of partial assignments of size at most w and satisfies the following conditions:

- The empty assignment ε belongs to $\mathcal{H}$.
- For every $\rho \in \mathcal{H}$, the assignment ρ does not falsify any clause of φ.
- If $\rho \in \mathcal{H}$ and $|\rho| < w$, then for every variable y_i on which ρ is undefined, there exists $\rho' \in \mathcal{H}$ such that $\rho \subseteq \rho'$, ρ' is defined on the variable y_i and $|\rho'| = |\rho| + 1$.

Atserias and Dalmau [5] showed that an unsatisfiable CNF formula φ admits a set of w-good assignments that is additionally *closed under taking subassignments* if and only if every resolution refutation of φ has width greater than w. For our purposes, we require a slightly more refined characterization capturing the absence of resolution refutations that are simultaneously of small width and small depth.

3.1 Characterization of Depth-Width Tradeoff

Consider a game between two players, Alice and Bob, defined with respect to an unsatisfiable CNF formula φ and two integer parameters w and h. Throughout the game, the players maintain a partial assignment to the variables of φ. Initially, the assignment is empty, and at all times its size is required to be at most w.

On each turn, Alice has two possible moves:

- If the size of the current assignment is less than w, Alice may ask Bob for the value of some variable x of φ. Bob then chooses and assigns a value to x.
- Alice may erase one variable from the domain of the current assignment.

Alice wins if the current assignment falsifies (contradicts) at least one clause of φ. Bob wins if he can answer at least h of Alice's questions without allowing Alice to win.

Now we define a winning strategy for Bob in this game.

Let $\mathcal{B}$ be a set of pairs (ρ, i) of a partial assignment ρ and an integer number i. We say that $\mathcal{B}$ is a (w, h)-winning strategy for φ if the following conditions hold:

- $(\varepsilon, 0) \in \mathcal{B}$, where ε is an empty assignment.
- If $(\rho, i) \in \mathcal{B}$, then $|\rho| \leq w$, $i \leq h$ and ρ doesn't falsify any clause of φ.
- If $(\rho, i) \in \mathcal{B}$ and $\rho' \subseteq \rho$, then $(\rho', i) \in \mathcal{B}$.
- If $(\rho, i) \in \mathcal{B}$, $|\rho| < w$, $i < h$, and $x \in \mathrm{Vars}(\varphi) \setminus \mathrm{Dom}(\rho)$, then there exists $a \in \{0, 1\}$ such that $(\rho \cup \{x := a\}, i + 1) \in \mathcal{B}$.

It is easy to see that if there exists a (w, h)-winning strategy $\mathcal{B}$ for φ, then Bob wins the game. Indeed, for every $i \in \{0, 1, \ldots, h\}$, Bob can guarantee that after his i-th response the current partial assignment ρ satisfies $(\rho, i) \in \mathcal{B}$. Consequently, ρ does not falsify any clause of φ, and Alice cannot win within h rounds.

Lemma 3.2 ([22])**.** Let $w \geq 0$ and $h \geq 0$ be some integers; and let φ be an unsatisfiable CNF formula such that φ doesn't have a resolution refutation of width at most w and simultaneously with depth at most h. Then there exists a (w, h)-winning strategy for φ.

Lemma 3.3. Let $\mathcal{B}$ be a (w, h)-winning strategy for φ, and define

$$\mathcal{H} := \{\tau \mid (\tau, |\tau|) \in \mathcal{B}\}.$$

Then $\mathcal{H}$ is a $\min\{w, h\}$-good set of assignments for φ.

Proof. Since $(\varepsilon, 0) \in \mathcal{B}$, it follows that $\varepsilon \in \mathcal{H}$.

For any $\rho \in \mathcal{H}$, we have $(\rho, |\rho|) \in \mathcal{B}$, and hence ρ does not falsify any clause of φ.

Now let $\rho \in \mathcal{H}$ satisfy $|\rho| < \min\{w, h\}$, and let y_i be a variable of φ on which ρ is not defined. Since $(\rho, |\rho|) \in \mathcal{B}$ and $|\rho| < \min\{w, h\}$, by the definition of a (w, h)-winning strategy there exists a value $a \in \{0, 1\}$ such that

$$(\rho \cup \{y_i := a\}, |\rho| + 1) \in \mathcal{B}.$$

Consequently, $\rho \cup \{y_i := a\} \in \mathcal{H}$. □

4 Applying Affine Restrictions to Parity Decision DAGs

In this section, we study applications of affine restrictions to parity decision DAGs and introduce the formal definition of applying an affine restriction. In the next subsection, we consider *structured affine restrictions* that can be applied to parity decision DAGs refuting lifted formulas $\varphi \circ g$. We show that, after such an application, the resulting graph is again a parity decision DAG refuting a lifted formula $\varphi' \circ g$, where φ' is obtained from φ by applying a partial assignment. Finally, in Subsection 4.2, we present a sufficient condition under which a linear system can be satisfied by a structured affine restriction of small size.

Let G be a parity decision DAG and ρ be an affine restriction. We define a parity decision DAG $G|_\rho$ as follows. First, we apply ρ to all linear systems labeling the nodes of G and to all linear equations labeling its edges. We then remove all edges whose labels become unsatisfiable under this substitution. Next, we delete all nodes whose associated linear systems become unsatisfiable, and subsequently remove all nodes that are no longer reachable from the source.

After these steps, some nodes may have an out-degree one. In this case, we merge such a node with its unique child and label the resulting node with the linear system of the child.

4.1 Affine Restrictions for Parity Decision DAGs Refuting Lifted Formulas

Proposition 4.1. Suppose a parity decision DAG G refutes a formula φ, and let ρ be an affine restriction that decomposes as a disjoint union $\rho = \rho_1 \cup \rho_2$, where ρ_1 is a partial assignment (i.e., an assignment that sets certain variables to 0 or 1). Assume that every clause of $\varphi|_{\rho_1}$ contains some clause of ψ, and that ψ does not involve any variables in the support of ρ_2. Then the restricted DAG $G|_\rho$ refutes ψ.

Proof. It is straightforward that $G|_{\rho_1}$ refutes $\varphi|_{\rho_1}$. Since every clause of $\varphi|_{\rho_1}$ contains some clause of ψ, it follows that $G|_{\rho_1}$ also refutes ψ. Finally, we have $G|_\rho = (G|_{\rho_1})|_{\rho_2}$, which refutes $\psi|_{\rho_2}$. But $\psi|_{\rho_2} = \psi$, since ψ does not involve any variables from the support of ρ_2. □

Proposition 4.2. Consider the lifted formula $\varphi \circ g$ and $i \in [m]$. Assume that a partial assignment ρ is supported on the subset of the ith block of lifted variables and assigns values to variables in such a way that, regardless of the values of the remaining unassigned lifted variables, ρ induces a partial assignment ρ' on the variable y_i. Then the formula $(\varphi \circ g)|_\rho$ is semantically equivalent to the formula $\varphi|_{\rho'} \circ g$ and, moreover, every clause of the former formula contains a clause of the latter formula.

Proof. Since lifting is defined clause-wise, it suffices to prove the statement for an arbitrary clause C of φ. If the clause C does not contain y_i, then neither ρ affects $C \circ g$ nor ρ' affects C.

So, we can assume that $C = y_i^a \vee D$.

Then $C \circ g$ is the standard CNF representation of the disjunction $g^a(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell}) \vee (D \circ g)$.

We consider two cases.

Case 1: ρ' satisfies y_i^a. Then, the Boolean function $g^a(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell})|_\rho$ is identically true, and hence every clause of the CNF formula $g^a(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell})$ is satisfied by ρ. Therefore, $C \circ g$ is satisfied by ρ. In this case, both formulas $(C \circ g)|_\rho$ and $C|_{\rho'} \circ g$ are identically true and thus contain no clauses in their CNF representations.

Case 2: ρ' falsifies y_i^a. Hence, the formula $g^a(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell})|_\rho$ is unsatisfiable.

Observe that $(C \circ g)|_\rho$ is the standard CNF representation of the disjunction $g^a(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell})|_\rho \vee (D \circ g)|_\rho$. Since ρ assigns variables only from the ith block, we have $(D \circ g)|_\rho = D \circ g$. So, $(C \circ g)|_\rho$ coincides with $g^a(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell})|_\rho \vee (D \circ g)$ and on the other hand, $C|_{\rho'} \circ g$ coincides with $D \circ g$.

As $g^a(x_{i,1}, x_{i,2}, \ldots, x_{i,\ell})|_\rho$ is unsatisfiable, it follows that $(C \circ g)|_\rho$ and $C|_{\rho'} \circ g$ are semantically equivalent. Moreover, every clause of the former formula contains a clause of the latter formula. □

Definition 4.3 ($\mathcal{H}$-good affine restriction)**.** Let $\mathcal{H}$ be a set of w-good assignments for an unsatisfiable CNF formula φ. An affine restriction σ on variables of $\varphi \circ g$ is called *$\mathcal{H}$-good* if it satisfies the following properties:

- σ is block-respecting and has block-size at most w.
- For every block, σ either fixes all variables to constants, or fixes all but one variable. In the latter case, the assignment to the fixed variables induces an assignment to the unlifted variable corresponding to that block, independent of the value of the remaining variable. The last variable in the block is restricted by σ to an affine function of the lifted variables from the blocks not touched by σ.
- The assignment induced by σ on the unlifted variables belongs to $\mathcal{H}$.

Propositions 4.2 and 4.1 imply the following lemma.

Lemma 4.4. Let $\mathcal{H}$ be a set of w-good assignments for an unsatisfiable CNF φ. Let σ be an $\mathcal{H}$-good affine restriction that induces a partial assignment $\rho \in \mathcal{H}$ on the unlifted variables. Let G be a parity decision DAG refuting $\varphi \circ g$, and let v be a node of G labeled with a linear system Φ_v such that σ satisfies Φ_v. Denote by G_v the subgraph of G consisting of all nodes reachable from v. Then the restricted DAG $G_v|_\sigma$ refutes $\varphi|_\rho \circ g$.

Proof. By definition, G_v is a parity decision DAG. Before applying σ, we cannot claim that G_v refutes $\varphi \circ g$, because the linear system at its source may be non-empty. However, after applying σ, the source of $G_v|_\sigma$ contains an empty (constant-true) linear system.

Let $\sigma = \sigma_1 \cup \sigma_2$ be a partition of σ, where σ_1 denotes the constant assignment (mapping variables to $\{0,1\}$) and σ_2 denotes the affine restriction defined on the rest of the domain. By the conditions of the lemma, σ_1 induces a partial assignment ρ on unlifted variables, independent of the values of the remaining variables in these blocks.

By applying Proposition 4.2 to all blocks on which σ_1 is defined, we obtain that the formula $(\varphi \circ g)|_{\sigma_1}$ is semantically equivalent to $\varphi|_\rho \circ g$ and every clause of the former formula contains a clause of the latter formula. Then by Proposition 4.1, the restricted DAG $G_v|_\sigma$ refutes $\varphi|_\rho \circ g$. □

4.2 Satisfying a Linear System by an $\mathcal{H}$-Good restriction

Since we consider only 1-stifling gadgets, for each lifted block it is possible to fix all but one of the lifted variables, thereby uniquely determining corresponding value of the unlifted variable. We exploit this property in the following lemma.

Lemma 4.5. Let g be a 1-stifling gadget and $\mathcal{H}$ be a set of w-good assignments for an unsatisfiable CNF φ. Suppose that a linear system Φ over lifted variables has a solution ρ such that the induced assignment $\text{induced}_g(\rho)$, when restricted to the variables with indices in $\text{Cl}(L(\Phi))$, belongs to $\mathcal{H}$. Assume further that $\dim(L(\Phi)) \leq w$. Then Φ can be satisfied by an $\mathcal{H}$-good restriction of block-size at most $\dim(L(\Phi))$.

Proof. Let σ_1 be the restriction of ρ to the lifted variables from the blocks in $\text{Cl}(L(\Phi))$. By assumption, σ_1 induces an assignment $\tau \in \mathcal{H}$ on the corresponding unlifted variables.

The restricted system $\Phi|_{\sigma_1}$ has a safe set of linear forms and is satisfiable. Hence, $\Phi|_{\sigma_1}$ can be equivalently rewritten in the form

$$\bigwedge_{s=1}^{r} x_{i_s,j_s} = a_s,$$

where the indices $i_1, i_2, \ldots, i_r$ are pairwise distinct, each a_s is an affine function that does not depend on the variables $x_{i_1,j_1}, x_{i_2,j_2}, \ldots, x_{i_r,j_r}$, and r is the rank of the system $\Phi|_{\sigma_1}$.

Observe that

$$r = \dim\langle L(\Phi)[\setminus \text{Cl}(L(\Phi))]\rangle.$$

By Lemma 2.9,

$$\dim(L(\Phi)) \geq |\,\text{Cl}(L(\Phi))| + \dim\langle L(\Phi)[\setminus \text{Cl}(L(\Phi))]\rangle,$$

and therefore $r + |\,\text{Cl}(L(\Phi))| \leq w$.

We now construct an $\mathcal{H}$-good restriction σ satisfying Φ. Extend σ_1 as follows. For each $s \in [r]$, assign $x_{i_s,j_s} := a_s$. Next, extend the induced assignment τ to the unlifted variable y_{i_s} so that τ remains in $\mathcal{H}$, and assign values to all remaining lifted variables in block i_s so as to force the corresponding unlifted block variables to agree with this extension of τ. Such an assignment is possible because g is 1-stifling.

The resulting restriction σ satisfies Φ, and its block-size is

$$|\,\text{Cl}(L(\Phi))| + \dim\langle L(\Phi)[\setminus \text{Cl}(L(\Phi))]\rangle,$$

which is at most $\dim(L(\Phi))$. □

5 Distribution on Paths

In this section, we present our main technical contribution: the construction of a random path.

Throughout this section, let φ be an unsatisfiable CNF formula over variables $Y = \{y_i \mid i \in [m]\}$, and let $g : \{0,1\}^\ell \to \{0,1\}$ be a gadget. We write $X = \{x_{i,j} \mid i \in [m], , j \in [\ell]\}$ for the set of lifted variables, that is, the variables of the lifted formula $\varphi \circ g$.

Let G be a parity decision DAG refuting $\varphi \circ g$. Consider the set of paths $\mathcal{P}(G)$ in G that start at the source. For a path $p \in \mathcal{P}(G)$, denote by Ψ_p the linear system consisting of linear equations written on the edges of p.

Theorem 5.1. Assume that $g : \{0,1\}^\ell \to \{0,1\}$ is a 1-stifling gadget. Let $\mathcal{H}$ be a set of w-good partial assignments for φ, where w is a positive integer.

Consider a parity decision DAG G refuting $\varphi \circ g$. There exists a distribution $\mathcal{D}$ supported on the paths, $\mathcal{P}(G)$, from the source of G with the following properties:

(1) With probability at least $1/4$, the length of a path $p \sim \mathcal{D}$ is at least $(w-1)/6$.
(2) For every p from the support of $\mathcal{D}$, a linear system Ψ_p has a solution τ and a set $J \subseteq [m]$ such that the restriction of $\text{induced}_g(\tau)$ to variables from blocks J belongs to $\mathcal{H}$ and $\text{Cl}(L(\Psi_p)) \subseteq J$.
(3) For every linear system Φ over variables X, $\Pr_{p\sim\mathcal{D}}[\Psi_p \models \Phi] \leq \exp\left(-\frac{1}{2^{\ell+1}\ell} \cdot \dim(L(\Phi))\right)$.

The rest of this section is devoted to the proof of Theorem 5.1.

5.1 Construction of the Distribution $\mathcal{D}$

We define the distribution $\mathcal{D}$ by specifying a randomized process that generates a path $p \sim \mathcal{D}$. The path p is constructed iteratively, where each step extends the path built so far. Initially, p has length 0 and consists of a single node—the source of G.

Along with the path, we maintain an $\mathcal{H}$-good affine restriction ρ, which is initially empty. Throughout the process, I denotes the set of blocks on which ρ is defined, and σ denotes the partial assignment to the unlifted variables induced by ρ. Since ρ is $\mathcal{H}$-good, we have $\sigma \in \mathcal{H}$.

We will also view affine restrictions as systems of linear equations. Specifically, an affine restriction $\{x_1 := h_1, x_2 := h_2, \ldots, x_k := h_k\}$ corresponds to the linear system $\bigwedge_{i=1}^{k}(x_i + h_i = 0)$.

Invariants. We maintain the following invariants throughout the construction:

- ρ is an $\mathcal{H}$-good affine restriction defined on the set of blocks I, and ρ induces the partial assignment σ to the unlifted variables. Recall that this means:
  - ρ is defined on the lifted variables belonging to the blocks in I;
  - for every $i \in I$, ρ either assigns 0/1-constants to all lifted variables of the ith block, or assigns constants to all but one lifted variable. In the latter case, regardless of the value of the remaining variable, the value of the unlifted variable is already determined by the gadget g;
  - ρ induces σ on the unlifted variables from the blocks in I, and $\sigma \in \mathcal{H}$.
- The length of p is at most $w/2$, and $|I| \leq w/2$;
- ρ, viewed as a linear system, semantically implies Ψ_p.

Variables in the domain of ρ that are not assigned constants by ρ are called *principal variables*. Since ρ is $\mathcal{H}$-good, each block in I contains at most one principal variable.

It is immediate that these invariants hold initially. In Claim 5.3 below, we will establish that as long as the invariants are preserved, the endpoint of the path p cannot be a sink of G.

Sampling algorithm. While $|p| < \lfloor w/2 \rfloor$ and $|I| \leq \lfloor w/2 \rfloor$, we attempt to extend the path p. Once either $|p|$ or $|I|$ reaches $\lfloor w/2 \rfloor$, the process terminates and outputs the path p.

Path prolongation. Let f be the linear form queried at the endpoint of the current path p, and consider the restriction $f|_\rho$.

- While $f|_\rho$ is not a constant and $|I| < \lfloor w/2 \rfloor$, perform the following steps:
  - *(1)* Choose a variable $x_{i,j}$ appearing in $f|_\rho$.
  - *(2)* Let $\alpha \in \{0,1\}$ be such that $\sigma \cup \{y_i := \alpha\} \in \mathcal{H}$.
  - *(3)* Toss a fair coin $c \in \{0,1\}$.
  - *(4)* **Branch A: sample the ith block.** If $c = 1$:
    - *(1)* Sample β uniformly at random from $g^{-1}(\alpha)$; recall that $g^{-1}(\alpha) \subseteq \{0,1\}^\ell$.
    - *(2)* Extend ρ on block i as follows:
      - Let π denote the partial assignment $\{x_{i,1} := \beta_1, \ldots, x_{i,\ell} := \beta_\ell\}$;
      - update $\rho := \rho|_\pi$, which substitutes the values of π into the right-hand sides of the assignments in ρ;
      - set $\rho := \rho \cup \pi$.
  - *(5)* **Branch B: assign a principal variable.** If $c = 0$:
    - *(1)* Consider a partial assignment π that sets all lifted variables of the ith block except $x_{i,j}$ so that the value of the gadget g on block i equals α, regardless of the value of $x_{i,j}$. Such an assignment exists since g is 1-stifling.
    - *(2)* Toss a fair coin $b \in \{0,1\}$.
    - *(3)* Extend ρ on block i as follows:
      - update $\rho := \rho|_\pi$, substituting the values of π into the right-hand sides of the assignments in ρ;
      - update $\rho := \rho|_{x_{i,j} := f|_{\rho \cup \pi} + x_{i,j} + b}$, substituting the new value of $x_{i,j}$ into the right-hand sides of the assignments in ρ;
      - set $\rho := \rho \cup \pi \cup \{x_{i,j} := f|_{\rho \cup \pi} + x_{i,j} + b\}$.
- If $f|_\rho$ is a constant, extend the path p along the edge labeled by the equation $f = f|_\rho$.
- If $|I| = \lfloor w/2 \rfloor$ or $|p| = \lfloor w/2 \rfloor$, terminate the process and return p.

Invariant verification. Observe that in both branches A and B, the restriction ρ remains $\mathcal{H}$-good. It is also immediate that $p \leq w/2$ and $|I| \leq w/2$ is preserved throughout the process.

Whenever the path p is extended by an edge labeled with an equation $f = a$, this equation is semantically implied by the current restriction ρ. Since each update to ρ only strengthens the linear system—so that the updated system semantically implies the previous one, it follows that at every stage the current restriction ρ semantically implies Ψ_p. Hence, all invariants are preserved.

Algorithm correctness.

Claim 5.2. Let π be a block-respecting partial assignment defined on the lifted variables of the blocks in $[m] \setminus I$. Then the linear system $\rho \wedge \pi$ is satisfiable.

Proof. Without loss of generality, assume that π is defined on all lifted variables from the blocks in $[m] \setminus I$. The system $\rho \wedge \pi$ is satisfiable because it suffices to assign values to the principal variables consistently with the values of the linear forms to which they are equated by ρ, as determined by π. □

Claim 5.3. If the invariants are satisfied, then the endpoint of the path p cannot be a sink of G.

Proof. It suffices to show that for every clause of $\varphi \circ g$, the system Ψ_p admits a solution that satisfies this clause. For this it is enough to prove that for any clause C of φ, there exists a solution to the linear system ρ that satisfies $C \circ g$.

Since $\sigma \in \mathcal{H}$, the partial assignment σ does not falsify C. Therefore, there exists a full assignment σ' to the unlifted variables that satisfies C. Now consider a block-respecting partial assignment π defined on the lifted variables of the blocks in $[m] \setminus I$ such that, on the unlifted variables, π induces the values prescribed by σ' on the corresponding blocks.

By Claim 5.2, the linear system $\rho \wedge \pi$ is satisfiable. By construction of the lifting, any solution to $\rho \wedge \pi$ satisfies $C \circ g$. This concludes the proof. □

Proof of Condition (1) in Theorem 5.1.

Claim 5.4. With probability at least $1/4$, the returned path has length at least $(w-1)/6$.

Proof. The process terminates once either $|p| = \lfloor w/2 \rfloor$ or $|I| = \lfloor w/2 \rfloor$.

In the first case, we immediately have $|p| \geq \lfloor w/2 \rfloor \geq (w-1)/6$, and there is nothing to prove. We therefore focus on the second case, where $|I| = \lfloor w/2 \rfloor$ at termination.

Observe that every time Branch B is executed (that is, when the coin c lands on 0), the restriction $f|_\rho$ becomes a constant, and hence the path p is extended by one edge. Consequently, the length of p is at least the number of times the coin c takes the value 0.

Each iteration of the loop increases $|I|$ by one, so the total number of coin tosses is exactly $\lfloor w/2 \rfloor$. Since c is uniformly distributed in $\{0,1\}$, the expected number of times $c = 1$ is $\frac{1}{2}\lfloor w/2 \rfloor$.

By Markov's inequality, the probability that the number of ones is at least $\frac{4}{3} \cdot \frac{1}{2}\lfloor w/2 \rfloor = \frac{2}{3}\lfloor w/2 \rfloor$ is at most $3/4$. Therefore, with probability at least $1/4$, the number of ones is less than $\frac{2}{3}\lfloor w/2 \rfloor$, which implies that the number of zeros is greater than $\frac{1}{3}\lfloor w/2 \rfloor$.

Hence, with probability at least $1/4$, we have $|p| \geq \frac{1}{3}\lfloor w/2 \rfloor \geq (w-1)/6$ as claimed. □

Proof of Condition (2) in Theorem 5.1.

Claim 5.5. If the invariants are satisfied, then Ψ_p admits a solution τ and there exists a set $J \subseteq [m]$ such that the restriction of $\mathrm{induced}_g(\tau)$ to the variables from blocks in J belongs to $\mathcal{H}$ and $\mathrm{Cl}(L(\Psi_p)) \subseteq J$.

Proof. Since ρ semantically implies Ψ_p, it suffices to exhibit a solution of the linear system ρ with the stated properties.

Note that $|I \cup \mathrm{Cl}(L(\Psi_p))| \leq |I| + |p| \leq w$. Therefore, the partial assignment $\sigma \in \mathcal{H}$ can be extended to an assignment $\sigma' \in \mathcal{H}$ that is defined on the unlifted variables from the blocks in

$I \cup \mathrm{Cl}(L(\Psi_p))$. Let π be a block-respecting partial assignment defined on the lifted variables of the blocks in $\mathrm{Cl}(L(\Psi_p)) \setminus I$ such that $\mathrm{induced}_g(\pi)$ agrees with σ' on the unlifted variables from these blocks. Such an assignment π exists because g is 1-stifling.

By Claim 5.2, the linear system $\rho \wedge \pi$ is satisfiable. Any solution τ of this system satisfies Ψ_p and has the required property for the set $J := I \cup \mathrm{Cl}(L(\Psi_p))$. □

In the next subsection, we establish Condition (3) in Theorem 5.1.

5.2 Estimating the Probability of Semantic Implication

In this subsection, we prove the following lemma.

Lemma 5.6. For any linear system Φ over variables X, it holds that

$$\Pr[\rho \models \Phi] \leq \left(1 - 2^{-(\ell+1)}\right)^{\dim(L(\Phi))/\ell},$$

where ρ is the final value of ρ and the probability is taken over random bits used in the algorithm generating the distribution D.

Since by the invariants, $\rho \models \Psi_p$ during the execution of the algorithm generating D, this is also true for their final values. Hence, Lemma 5.6 implies that $\Pr_{p\sim D}[\Psi_p \models \Phi] \leq \Pr[\rho \models \Phi] \leq (1 - 2^{-(\ell+1)})^{\dim(L(\Phi))/\ell} \leq \exp\left(-\frac{1}{2^{\ell+1}\ell} \cdot \dim(L(\Phi)\right)$, thus, establishing Condition (3) in Theorem 5.1.

To prove Lemma 5.6, we carefully analyze what happens during the generation algorithm. Let us denote by V_ρ the linear space generated by linear forms $L(\rho)$. Denote by U the linear space generated by linear forms $L(\Phi)$. If $\rho \models \Phi$, then any equation of Φ is a linear combination of several equations of ρ, hence $U \subseteq V_\rho$. We look at how $V_\rho \cap U$ is changing during the execution of the generation algorithms. Initially ρ is an empty system, thus $V_\rho \cap U = \varnothing$.

Notice that ρ updates in the manner such that the updated ρ semantically implies the previous values. Thus, once we receive ρ that is not consistence with Φ, it will be inconsistent with Φ up to the end and $\rho \not\models \Phi$.

There are two different ways in which ρ can be modified. In Branch A, we substitute constants to all variables from the ith block, so $L(\rho)$ is increased by $\{x_{i,k} \mid k \in [\ell]\}$. In Branch B, we substitute constants to all variables from the ith block but $x_{i,j}$ and $L(\rho)$ is increased by $\{x_{i,k} \mid k \in [\ell]\setminus\{j\}\}\cup\{f\}$.

Lemma 5.7. Consider Step (2) of the algorithm (i.e., we have not yet tossed coin c). Assume that either $\dim(\langle V_\rho \cup \{x_{i,k} \mid k \in [\ell]\}\rangle \cap U)$ or $\dim(\langle V_\rho \cup \{x_{i,k} \mid k \in [\ell] \setminus \{j\}\} \cup \{f\}\rangle \cap U)$ is greater than $\dim(V_\rho \cap U)$. Then, with probability at least $2^{-(\ell+1)}$ after execution of the current iteration of the cycle ρ will be inconsistent with Φ.

Proof. **Case 1.** Adding linear forms $x_{i,k}$ for $k \in [\ell]$ to V_ρ increases its intersection with U:

Since the intersection $\langle V_\rho \cup \{x_{i,k} \mid k \in [\ell]\}\rangle \cap U$ is strictly greater than $V_\rho \cap U$, there is a subset $S \subseteq [\ell]$ and $v \in V$, such that $v + x_{i,S} \in U \setminus V$, where $x_{i,S}$ denotes $\sum_{j\in S} x_{i,j}$. Since $v + x_{i,S} \in U$, Φ semantically implies some fixed value $\gamma \in \{0, 1\}$ for linear form $v + x_{i,S} \in U$. Since $v \in V$, ρ implies some fixed value $\delta \in \{0, 1\}$ for linear form v.

With probability $\frac{1}{2}$, we toss $c = 1$, and the algorithm will go to Branch A. And to the lifted variables from the block i we substitute the random value from $g^{-1}(\alpha)$. Since g is 1-stifling, by Lemma 2.5, with probability at least $2^{-\ell}$ the value of the linear form $x_{i,S}$ will be different from

$\gamma + \delta$. Hence, after the Branch A, ρ and Φ will be inconsistent since they imply different values to the linear form $v + x_{i,S}$. And this happens with probability at least $\frac{1}{2} \cdot 2^{-\ell} = 2^{-(\ell+1)}$.

Case 2. Assume the condition of case 1 does not hold but adding linear forms $\{x_{i,k} \mid k \in [\ell] \setminus \{j\}\} \cup \{f\}\}$ to V_ρ increases its intersection with U.

Then there is a subset $S \subseteq [\ell] \setminus \{j\}$ and $v \in V$, such that $v + f + x_{i,S} \in U \setminus V$. Since $v + f + x_{i,S} \in U$, Φ semantically implies some fixed value $\gamma \in \{0,1\}$ for linear form $v + f + x_{i,S} \in U$. Since $v \in V$, ρ implies some fixed value $\delta \in \{0,1\}$ for linear form v.

With probability $\frac{1}{2}$, we toss $c = 0$, and the algorithm will go to Branch B. Then ρ will fix $x_{i,S}$ by some way and fix f uniformly at random. Hence, the probability $\frac{1}{2}$, ρ and Φ implies different values of the linear form $v + f + x_{i,S}$. Thus, totally with probability at least $\frac{1}{4} \geq 2^{-(\ell+1)}$, ρ and Φ will be inconsistent. □

Now we are ready to prove Lemma 5.6.

Proof of Lemma 5.6. Observe that after each iteration of the main loop (that is, after executing either Branch A or Branch B), the quantity $\dim(V_\rho \cap U)$ increases by at most ℓ. If the final value of Ψ_p semantically implies Φ, then so does the final value of ρ. In particular, this means that $\dim(V_\rho \cap U) = \dim(L(\Phi))$.

Consequently, there must be at least $\dim(L(\Phi))/\ell$ iterations of the main loop during which the intersection $V_\rho \cap U$ increases.

By Lemma 5.7, in each such iteration, with probability at least $2^{-(\ell+1)}$, the restriction ρ becomes inconsistent with Φ. Therefore, the probability that the final restriction ρ semantically implies Φ is at most

$$\left(1 - 2^{-(\ell+1)}\right)^{\dim(L(\Phi))/\ell}.$$

□

6 The Main Lifting Theorem

In this section, we prove our main lifting theorem and its corollaries.

Theorem 6.1. Assume that every resolution refutation of an unsatisfiable CNF formula ϕ has either width at least w, or depth at least h. Let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-stifling gadget, where ℓ is a constant. Then any Res($\oplus$) refutation of $\phi \circ g$ of size S has depth at least $\Omega\left(\min\{h, w^2/\log S\}\right)$.

As outlined in the introduction, Theorem 6.1 is proved by induction on the width w. We begin with the lemma that forms the core of the induction step.

Let $\mathcal{H}$ be a set of w-good assignments for φ. We say that $\mathcal{H}$ has one-time rollback property if for every block-respecting τ that is subassignment of some assignment from $\mathcal{H}$, there is a set of w-good assignments $\mathcal{H}_\tau$ for φ such that $\tau \in \mathcal{H}_\tau$.

For example, if $\mathcal{H}$ is closed under restrictions, then $\mathcal{H}$ has one time rollback property and $\mathcal{H}_\tau = \mathcal{H}$.

Lemma 6.2. Let φ be an unsatisfiable CNF formula, and let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-stifling gadget, where ℓ is a constant. Let $\mathcal{H}$ be a set of w-good partial assignments for φ. Suppose that G is a parity decision DAG refuting $\varphi \circ g$ of size S. Then there exists an absolute constant $c > 0$ and a node v of G such that:

– there is a path from the source of G to v of length at least $(w-1)/6$;
– if $\mathcal{H}$ has the one-time rollback property and $w > c \log S$, then there exists an affine restriction ρ such that $G_v|_\rho$ is a parity decision DAG refuting $\varphi|_\sigma \circ g$, where:
  – G_v denotes the subgraph of G consisting of all nodes reachable from v;
  – σ is a partial assignment to the unlifted variables of size at most $c \log S$, and $\sigma \in \mathcal{H}_\tau$ for some subassignment τ of an assignment from $\mathcal{H}$.

Proof. By Theorem 5.1, there exists a distribution D supported on paths from the source of G such that:

(1) With probability at least $1/4$, a path $p \sim D$ has length at least $(w-1)/6$.
(2) For every p from the support of $\mathcal{D}$, a linear system Ψ_p has a solution π and a set $J \subseteq [m]$ such that the restriction of $\text{induced}_g(\pi)$ to variables from blocks J belongs to $\mathcal{H}$ and $\text{Cl}(L(\Psi_p)) \subseteq J$.
(3) For every linear system Φ over variables X, $\Pr_{p\sim D}[\Psi_p \models \Phi] \leq 2^{-\Omega(\dim(L(\Phi)))}$.

For a path p starting at the source of G, let v_p denote its endpoint and let Φ_{v_p} be the linear system labeling v_p. Let E be the set of all systems Φ_{v_p} corresponding to paths p in the support of D whose length is at least $(w-1)/6$. By construction, $\Pr_{p\sim D}[\Phi_{v_p} \in E] \geq \frac{1}{4}$.

On the other hand, for every $\Phi \in E$,

$$\Pr_{p\sim D}[\Phi_{v_p} = \Phi] \leq \Pr_{p\sim D}[\Psi_p \models \Phi] \leq 2^{-\Omega(\dim(L(\Phi)))},$$

where the first inequality follows from Lemma 2.2. Since $|E| \leq S$, there exists a path p_0 of length at least $(w-1)/6$ such that $\dim(\Phi_{v_{p_0}}) \leq O(\log(4S)) = O(\log S)$. Let $c > 0$ be an absolute constant such that $\dim(\Phi_{v_{p_0}}) \leq c \log S$.

By Lemma 2.2, we have $\Psi_{p_0} \models \Phi_{v_{p_0}}$. Since Ψ_{p_0} is satisfiable, it follows that $L(\Phi_{v_{p_0}}) \subseteq \langle L(\Psi_{p_0})\rangle$, and by the properties of the closure (Lemma 2.9),

$$\text{Cl}(L(\Phi_{v_{p_0}})) \subseteq \text{Cl}(L(\Psi_{p_0})).$$

Let π be a solution to Ψ_{p_0} witnessing the second statement of Theorem 5.1. Then π also satisfies $\Phi_{v_{p_0}}$. Let τ denote the restriction of $\text{induced}_g(\pi)$ to the variables from blocks in $\text{Cl}(L(\Phi_{v_{p_0}}))$. Then τ is a subassignment of some assignment in $\mathcal{H}$. Since $\mathcal{H}$ has the one-time rollback property, there exists a family $\mathcal{H}_\tau$ of w-good partial assignments for φ that contains τ.

By Lemma 4.5, the system $\Phi_{v_{p_0}}$ can be satisfied by an $\mathcal{H}_\tau$-good affine restriction ρ of block-size at most $O(\log(S))$. Let $\sigma \in \mathcal{H}_\tau$ be an assignment induced by ρ on the unlifted variables; the size of σ is at most $O(\log S)$. Consider the subgraph $G_{v_{p_0}}$ of G consisting of all nodes reachable from v_{p_0}. By Lemma 4.4, applying ρ to $G_{v_{p_0}}$ yields a parity decision DAG refuting $\varphi|_\sigma \circ g$. □

Proof of Theorem 6.1. By Lemma 3.2, there exists a (w,h)-winning strategy for ϕ.

Define $d(w,h,S)$ as the minimum depth of a parity decision DAG of size at most S refuting $\varphi \circ g$, where the minimum ranges over all CNF formulas φ possessing a (w,h)-winning strategy. Our goal is to show that $d(w,h,S) \geq \Omega\left(\min\{h, w^2/\log S\}\right)$. Assume that this minimum is attained by a formula φ, and fix a (w,h)-winning strategy $\mathcal{B}$ for φ.

Consider $\mathcal{H} := \{\tau \mid (\tau, |\tau|) \in \mathcal{B}\}$. By Lemma 3.3, $\mathcal{H}$ is a $\min\{w,h\}$-good set of assignments for φ.

If $h < 2w$, then the theorem follows from Lemma 6.2. So we assume that $h \geq 2w$. In this case, $\mathcal{H}$ has the one-time rollback property. Indeed, consider some $\tau \in \mathcal{H}$, we know that $(\tau, |\tau|) \in \mathcal{B}$. If

$\sigma \subseteq \tau$, then $(\sigma, |\tau|) \in \mathcal{B}$. Let us define $\mathcal{H}_\sigma := \{\pi \mid \exists i \leq |\tau| + |\pi| : (\pi, i) \in \mathcal{B}\}$. It is easy to see that $\mathcal{H}_\sigma$ is a set of w-good assignments for φ and that $\sigma \in \mathcal{H}_\sigma$.

Consider a parity decision DAG G refuting $\varphi \circ g$, let S be an upper bound on the size of G. By Lemma 6.2 applied to $\mathcal{H}$ there is a node v of G such that there is an absolute constant $c > 0$ and a path from the source to v of length at least $(w-1)/6$ (we denote by G_v a subgraph of G that contains all nodes that are reachable from v) and there is an affine restriction ρ of block-size at most $c \log S$ and a partial assignment σ from $\mathcal{H}_\tau$ for some $\tau \in \mathcal{H}$ of size at most $c \log S$ such that $G_v|_\rho$ is a parity decision DAG refuting $(\varphi|_\sigma) \circ g$.

Since the block-size of σ is at most $c \log S$, there exists a $(h - w - \lceil c \log S \rceil, w - \lceil c \log S \rceil)$-winning strategy $\mathcal{B}'$ for $\varphi|_\sigma$; for example if $(\sigma, k) \in \mathcal{B}$, we can choose

$$\mathcal{B}' = \{(\pi \setminus \sigma, j) \mid (\pi, j + k) \in \mathcal{B}, \sigma \subseteq \pi\}.$$

Therefore,

$$d(w, h, S) \geq \frac{w-1}{6} + d(w - \lceil c \log S \rceil, h - w - \lceil c \log S \rceil, S).$$

Unrolling the recurrence yields

$$d(h, w, S) \geq \Omega(\min\{h, w^2 / \log S\}).$$

□

We apply this theorem to the supercritical size-depth tradeoff for resolution by Buss and Thapen [10]. It is convenient to use the corollary of this tradeoff presented in [22].

Theorem 6.3 ([10, 22])**.** There exists a family of unsatisfiable CNF formulas $\{\Psi_n\}_{n=1}^\infty$ such that

- Ψ_n contains n variables;
- the width of Ψ_n is $O(\log n)$ and, moreover, Ψ_n has a resolution refutation of size $\mathrm{poly}(n)$ and of width $O(\log n)$;
- any resolution refutation of Ψ_n of width at most $n/40 \log n$ has depth greater than $n^2/400 \log^2 n$.

Theorem 6.4. There is a family of formulas ψ_n from n variables of size $\mathrm{poly}(n)$ such that ψ_n has polynomial size resolution refutation, and for any $S > 0$, any Res($\oplus$) refutation of size S has depth at least $\Omega(n^2/(\log^2 n \log S))$.

Proof. Let Ψ_n be a formula from Theorem 6.3. Let $g = \mathrm{MAJ}_3$; it is a 1-stifling gadget. Then $\Psi_n \circ g$ is a formula in $O(\log n)$-CNF of size $\mathrm{poly}(n)$ and it has a resolution refutation of size $\mathrm{poly}(n)$. By Theorem 6.1, any Res($\oplus$) refutation of $\Psi_n \circ g$ of size at most S has depth at least $\Omega(n^2/(\log^2 n \log S))$. □

Corollary 6.5. If $d(n) = o(n^2/\log^4 n)$, then Depth-$d(n)$ Res($\oplus$) does not polynomially simulates resolution.

Proof. Consider the formula ψ_n from Theorem 6.4. Applying Theorem 6.4 with $S = 2^{\log^2 n}$, we obtain that any Res($\oplus$) refutation of ψ_n of depth $d(n)$ must have size at least $2^{\log^2 n}$, which is superpolynomial in n. □

7 A Quadratic Size Lower Bound for Unrestricted Res($\oplus$)

In this section, we present a quadratic size lower bound for Res($\oplus$). We begin with a linear-algebraic lemma, next recall the necessary previous results, and then conclude with the proof of the main theorem.

Lemma 7.1. Let $V_1, V_2, \ldots, V_s$ be linear subspaces of $\mathbb{F}^n$, where $\mathbb{F}$ is a field. Assume that $V_1 = \{0\}$ and that for every $i \in [s-1]$ there exists a vector $x_i \in \mathbb{F}^n$ such that $V_{i+1} \subseteq \langle V_i \cup \{x_i\}\rangle$. Then there exists a nonzero vector x such that $\big|\{i \in [s] \mid x \in V_i\}\big| \geq \dim V_s$, that is, x belongs to at least $\dim V_s$ of the subspaces V_i.

Proof. We prove the lemma by induction on $k := \dim V_s$. The base cases $k = 0$ and $k = 1$ are immediate. Assume that $k > 1$.

For each $i \in [s]$, define $V'_i := V_i \cap V_s$. Note that

$$V'_{i+1} = V_{i+1} \cap V_s \subseteq \langle V_i \cup \{x_i\}\rangle \cap V_s.$$

Notice that,

$$\dim\big(\langle V_i \cup \{x_i\}\rangle \cap V_s\big) \leq \dim(V_i \cap V_s) + 1 = \dim(V'_i) + 1.$$

Therefore, there exists a vector $x'_i \in \mathbb{F}^n$ such that

$$\langle V_i \cup \{x_i\}\rangle \cap V_s \subseteq \langle (V_i \cap V_s) \cup \{x'_i\}\rangle = \langle V'_i \cup \{x'_i\}\rangle,$$

and hence $V'_{i+1} \subseteq \langle V'_i \cup \{x'_i\}\rangle$ for every $i \in [s-1]$.

Since $V'_1 = \{0\}$, $\dim(V'_s) = k$, and for every $i \in [s-1]$ we have $\dim(V'_{i+1}) \leq \dim(V'_i) + 1$, there exists $j \in [s-1]$ such that $\dim(V'_j) = k - 1$.

Applying the induction hypothesis to the sequence $V'_1, V'_2, \ldots, V'_j$, we obtain a nonzero vector x that belongs to at least $k-1$ of these subspaces. Since V_s contains every V'_i, the vector x also belongs to V_s. Hence x belongs to at least k of the original subspaces $V_1, \ldots, V_s$ (namely, to at least $k-1$ among $V_1, \ldots, V_j$, and also to V_s), completing the induction step. □

We will use the following result of Alekseev and Itsykson [4], which shows that resolution width lifts to width in Res($\oplus$).

Theorem 7.2 ([4])**.** Let φ be an unsatisfiable CNF formula such that every resolution refutation of φ has width at least w. Let g be a 1-stifling gadget. Then every Res($\oplus$) refutation of $\varphi \circ g$ has width at least w.

We will also need the following basic fact about resolution width.

Lemma 7.3 ([8])**.** Let φ be an unsatisfiable CNF formula of resolution width w, and let x be a variable of φ. Then there exists a value $a \in \{0,1\}$ such that the restricted formula $\varphi|_{x:=a}$ has resolution width at least $w-1$.

Theorem 7.4. Let $\varphi(y_1, y_2, \ldots, y_m)$ be an unsatisfiable CNF formula such that every resolution refutation of φ has width at least w. Let $g : \{0,1\}^\ell \to \{0,1\}$ be a 1-stifling gadget. Then any Res($\oplus$) refutation of $\varphi \circ g$ has size at least $\frac{w(w+1)}{2}$.

Proof. We prove the theorem by induction on w. The base case $w = 1$ is trivial.

Let G be a parity decision DAG refuting $\varphi \circ g$ of minimum size. By Theorem 7.2, G contains a node v such that the linear system Φ_v labeling this node has rank at least w. Fix a path $v_1, \ldots, v_s = v$ from the source v_1 to v. Since G has minimum size, we may assume that for all $i \in [s]$ the system Φ_{v_i} is satisfiable, and that for all $i \in [s-1]$ the equation labeling the edge (v_i, v_{i+1}) is consistent with Φ_{v_i}.

For each $i \in [s]$, let V_i denote the vector space of linear equations spanned by the equations of Φ_{v_i}. We have $V_1 = \{0 = 0\}$ and $\dim(V_s) = \mathrm{rk}(\Phi_v) \geq w$. Moreover, for every $i \in [s-1]$, the space V_{i+1} is contained in the span of V_i and the equation labeling the edge (v_i, v_{i+1}). By Lemma 7.1, there exists a nontrivial equation $f = a$ that belongs to at least w of the spaces V_i, for $i \in [s]$.

Since all systems Φ_{v_i} are satisfiable, $f = a$ is different from $0 = 1$. In particular, the linear form f contains at least one variable x_{i_0,j_0}.

By Lemma 7.3, there exists $a \in \{0,1\}$ such that the restricted formula $\varphi|_{y_{i_0}:=a}$ has resolution width at least $w-1$. Let ρ_1 be a partial assignment that assigns Boolean values to all lifted variables from the i_0th block except x_{i_0,j_0}, in such a way that these assignments force $y_{i_0} = a$ independently of the value of x_{i_0,j_0}. Such an assignment ρ_1 exists since g is 1-stifling. By Proposition 4.2, the restricted formula $(\varphi \circ g)|_{\rho_1}$ is semantically equivalent to $\varphi|_{y_{i_0}:=a} \circ g$, and every clause of the former contains a clause of the latter.

Next, define an affine restriction ρ_2 supported on the one variable

$$x_{i_0,j_0} := f|_{\rho_1} + x_{i_0,j_0} + a + 1,$$

and let $\rho := \rho_1 \cup \rho_2$. By Proposition 4.1, the restricted DAG $G|_\rho$ refutates $\varphi|_{y_{i_0}:=a} \circ g$.

Since the resolution width of $\varphi|_{y_{i_0}:=a}$ is at least $w-1$, the induction hypothesis implies that the size of $G|_\rho$ is at least $\frac{(w-1)w}{2}$. On the other hand, G contains at least w nodes whose linear systems are falsified by ρ, namely those whose associated spaces contain the equation $f = a$. Therefore, the size of $G|_\rho$ is smaller than the size of G by at least w nodes. It follows that the size of G is at least $\frac{w(w+1)}{2}$, completing the proof. □

References

[1] M. Ajtai. $\sum_1^1$-formulae on finite structures. *Annals of Pure and Applied Logic*, 24(1):1–48, 1983.

[2] Miklós Ajtai. The complexity of the pigeonhole principle. *Combinatorica*, 14(4):417–433, 1994.

[3] Yaroslav Alekseev and Nikita Gaevoy. New polynomial-depth Res(+) lower bounds. *Electron. Colloquium Comput. Complex.*, TR26-007, 2026.

[4] 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.

[5] Albert Atserias and Víctor Dalmau. A combinatorial characterization of resolution width. *J. Comput. Syst. Sci.*, 74(3):323–334, 2008.

[6] 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)*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14:1–14:17, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[7] Christopher Beck and Russell Impagliazzo. Strong ETH holds for regular resolution. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 487–494. ACM, 2013.

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

[9] Sreejata Kishor Bhattacharya and Arkadev Chattopadhyay. Exponential lower bounds on the size of reslin proofs of nearly quadratic depth. *Electron. Colloquium Comput. Complex.*, TR25-106, 2025.

[10] Sam Buss and Neil Thapen. A simple supercritical tradeoff between size and height in resolution. *Information Processing Letters*, 191:106589, 2026. The preprint is available at `https://eccc.weizmann.ac.il/report/2024/001`.

[11] Sam Buss and Neil Thapen. A simple supercritical tradeoff between size and height in resolution. *Inf. Process. Lett.*, 191:106589, 2026.

[12] Farzan Byramji and Russell Impagliazzo. Lifting to randomized parity decision trees. In Alina Ene and Eshan Chattopadhyay, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2025, Berkeley, CA, USA, August 11-13, 2025*, volume 353 of *LIPIcs*, pages 55:1–55:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025.

[13] Farzan Byramji and Russell Impagliazzo. Lower bounds for bit pigeonhole principles in bounded-depth resolution over parities. *CoRR*, abs/2511.20023, 2025.

[14] 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.

[15] Klim Efremenko, Michal Garlík, and Dmitry Itsykson. Lower bounds for regular resolution over parities. *SIAM J. Comput.*, 54(4):887–915, 2025.

[16] 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)*, volume 339 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 8:1–8:24, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[17] Klim Efremenko and Dmitry Itsykson. Strong ETH holds for bounded-depth resolution over parities. *Electron. Colloquium Comput. Complex.*, TR25-188, 2025.

[18] Merrick Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *22nd Annual Symposium on Foundations of Computer Science (sfcs 1981)*, pages 260 – 270, 1981.

[19] Ankit Garg, Mika Göös, Pritish Kamath, and Dmitry Sokolov. Monotone circuit lower bounds from resolution. *Theory Comput.*, 16:1–30, 2020.

[20] Michal Garlík and Leszek Aleksander Kolodziejczyk. Some subsystems of constant-depth frege with parity. *ACM Trans. Comput. Log.*, 19(4):29:1–29:34, 2018.

[21] Svyatoslav Gryaznov, Sergei Ovcharov, and Artur Riazanov. Resolution over linear equations: Combinatorial games for tree-like size and space. *ACM Trans. Comput. Theory*, 16(3):15:1–15:15, 2024.

[22] Dmitry Itsykson and Alexander Knop. Supercritical tradeoff between size and depth for resolution over parities. In Shubhangi Saraf, editor, *17th Innovations in Theoretical Computer Science Conference, ITCS 2026, Bocconi University, Milan, Italy, January 27-30, 2026*, volume 362 of *LIPIcs*, pages 81:1–81:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2026.

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

[24] 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.

[25] Dmitry Itsykson and Dmitry Sokolov. Resolution over linear equations modulo two. *Ann. Pure Appl. Log.*, 171(1), 2020.

[26] Erfan Khaniki. On proof complexity of resolution over polynomial calculus. *ACM Trans. Comput. Log.*, 23(3):16:1–16:24, 2022.

[27] Jan Krajíček. Randomized feasible interpolation and monotone circuits with a local oracle. *J. Math. Log.*, 18(2):1850012:1–1850012:27, 2018.

[28] Vladimir Podolskii and Alexander Shekhovtsov. Randomized lifting to semi-structured communication complexity via linear diversity. In Raghu Meka, editor, *16th Innovations in Theoretical Computer Science Conference, ITCS 2025, January 7-10, 2025, Columbia University, New York, NY, USA*, volume 325 of *LIPIcs*, pages 78:1–78:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025.

[29] A. A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mat. Zametki*, 41:598–607, 1987.

[30] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 77–82, 1987.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il