Electronic Colloquium on Computational Complexity, Revision 3 of Report No. 139 (2026)

Fine-Grained AC^0 Lower Bounds for k-OV, k-XOR, and k-SUM via Colored Subgraph Isomorphism

Haoxing Lin*
National University of Singapore
haoxingl@comp.nus.edu.sg

Abstract

We study the circuit cost of selecting k objects satisfying orthogonality, zero XOR, or modular zero SUM. Our main lower bounds in nonuniform AC^0 use a common framework of depth-zero projections from colored subgraph isomorphism. The projections use dimension, row count, or bit width $O(k \log n)$ and preserve circuit size, depth, and gate orientation. OV inputs satisfy pairwise distinctness within each array. Our direct XOR bounds require odd k, while SUM admits both parities. Stronger lower bounds for suitable source patterns transfer through the same framework to all three targets.

For every fixed depth d and sufficiently large fixed k, we prove unconditional lower bounds $n^{\Omega(k)}$ for OV and $(n/k)^{\Omega(k)}$ for XOR and SUM, with an absolute constant in the exponent independent of d and k; only the threshold on n may depend on (d, k). For $k \to \infty$ with $k = n^{o(1)}$, combining our projections with Beame's small-clique lower bound yields $n^{\Omega_d(\min\{\sqrt{k}, \log n\})}$ at every fixed depth d. We strengthen the exponent to $\Omega(k)$ at depth two in both orientations and at depth three with a top disjunction. The depth-three result allows unrestricted bottom fan-in and literal polarity; its proof bounds the probability that a planted copy forms a minterm of an arbitrary fixed CNF over a random background.

We show that a pattern-uniform strengthening of the Li–Razborov–Rossman lower bound would give $n^{\Omega_d(k)}$ in the same growing-k regime for depth three with a top conjunction and every fixed depth $d \geq 4$. This is our only conditional main result; no running-time hypothesis is used. Supplementary results cover general budgets (dimension, row count, bit width), an odd-to-even XOR lift under explicit arithmetic conditions, and upper bounds. At SUM bit-width $m = \Theta(k \log(en/k))$ with a sufficiently large constant factor, a top-disjunction depth-three upper bound of size $(n/k)^{O(k)}$ matches our $(n/k)^{\Omega(k)}$ lower bound up to constants in the exponent, for sufficiently large k and $n \geq k^C$ with a suitable constant $C > 1$.

*This work was supported by the National Research Foundation, Singapore, under its NRF Fellowship programme, award no. NRF-NRFF14-2022-0010.

ISSN 1433-8092

Contents

1 Introduction **3**
- 1.1 A common framework and its source bounds . . . 3
- 1.2 Main results . . . 4
- 1.3 Related work . . . 6

2 The Projection Framework and The Source Problem **7**
- 2.1 The source problem and shared patterns . . . 7
- 2.2 The circuit model and common projection construction . . . 8
- 2.3 Master projections and host substitutions . . . 11

3 Three Constructions **12**
- 3.1 k-Orthogonal Vectors . . . 13
- 3.2 Odd k-XOR . . . 16
- 3.3 k-SUM . . . 18

4 Unconditional Lower Bounds **21**
- 4.1 Fixed parameters . . . 21
- 4.2 A clique bound at every fixed depth . . . 22
- 4.3 Depth two: counting terms and clauses . . . 24
- 4.4 Depth three: minterms and a top disjunction . . . 27

5 Conditional Bounds and Open Problems **29**
- 5.1 A conditional consequence . . . 29
- 5.2 Open problems . . . 30

6 Acknowledgments **30**

7 AI Disclosure **30**

A Source Functions and Cited Lower Bounds **33**

B Pattern Constructions and Constants **38**

C Lower Bounds for Other Encoding Budgets **41**

D Supplementary Results for k-XOR **45**

E Upper Bounds and Comparison with Lower Bounds **51**

F Proofs for the Depth-Three Bounds **54**

1 Introduction

How much circuit size is needed to select k objects satisfying one joint relation? The three problems studied here express this witness-selection task in different algebras:

Definition 1.1 (The three target problems)**.**

(i) k-OV: For $n \le 2^D$, an input to $k\text{-OV}_{n,D,k}$ consists of k ordered arrays $U_1, \dots, U_k$, each containing n vectors in $\{0,1\}^D$; this is a knD-bit encoding. Write $U_i = (u_{i,1}, \dots, u_{i,n})$, and impose the promise that the vectors within each array are pairwise distinct. The arrays thus represent k sets of vectors, as in the standard set formulation of k-OV. The function accepts exactly when there are indices $j_1, \dots, j_k \in [n]$ such that $\prod_{i=1}^{k} u_{i,j_i}(t) = 0$ for every $t \in [D]$.

(ii) k-XOR: An input to $k\text{-XOR}_{n,m,k}$ is an indexed list $(a_1, \dots, a_n) \in (\mathbb{F}_2^m)^n$. The function accepts exactly when there is a set $S \in \binom{[n]}{k}$ such that $\bigoplus_{i \in S} a_i = 0^m$.

(iii) k-SUM: An input to $k\text{-SUM}_{n,m,k}$ is an indexed list $(z_1, \dots, z_n) \in (\mathbb{Z}_{2^m})^n$. The function accepts exactly when there is a set $S \in \binom{[n]}{k}$ such that $\sum_{i \in S} z_i \equiv 0 \pmod{2^m}$.

XOR and SUM use nm input bits: each XOR entry uses its m coordinates, and each SUM entry uses the m-bit binary encoding of its representative in $[0, 2^m)$. Orthogonality requires that no coordinate be one in every selected vector, XOR checks linear equations over $\mathbb{F}_2$, and modular SUM checks integer addition with carries. We study the dependence of the required circuit-size exponent on k at dimension, row count, or bit width $O(k \log n)$.

We give a common projection framework that transfers lower bounds from colored subgraph isomorphism to all three problems. In this source problem, the input graph (the *host graph*) has one color class of size N for each vertex of a prescribed pattern P. A witness chooses one vertex from each class so that every pattern edge is present between the corresponding chosen vertices; we call such a selection a *colored copy* of P. Equivalently, it selects one present host edge for each pattern edge, with all choices agreeing at shared endpoints.

For every fixed pattern P and circuit depth d, the lower bound of Li–Razborov–Rossman (LRR) [LRR17] holds for all sufficiently large N, but the fixed-pattern statement does not control how the required threshold grows with P. It therefore gives no uniform guarantee when the pattern grows with k. In this regime, we combine the same projections with source bounds that explicitly relate N to the pattern size.

Circuit model and notation. We use nonuniform circuits with unbounded-fan-in AND/OR gates and negations at input leaves. For fixed d, $\text{size}_d(f)$ is the minimum number of non-input gates in a circuit for f of depth at most d; for OV, correctness is required on all inputs satisfying the distinctness promise. We write $\Sigma_3 = \text{OR}\circ\text{AND}\circ\text{OR}$ and $\Pi_3 = \text{AND}\circ\text{OR}\circ\text{AND}$. Constants hidden by Ω are absolute unless a subscript specifies their dependence. Our use of "fine-grained" concerns the dependence of the circuit-size exponent on k. Conditional running-time bounds motivate that question [IP01, IPZ01, VW18, CW19] but do not by themselves constrain nonuniform circuits; they supply no hardness assumption here.

1.1 A common framework and its source bounds

Encoding colored copies in the three problems. To transfer lower bounds from the source problem, we encode colored copies as selections of k vectors in OV or k list entries in XOR and SUM. The constructed target instance has a witness exactly when the source contains a colored

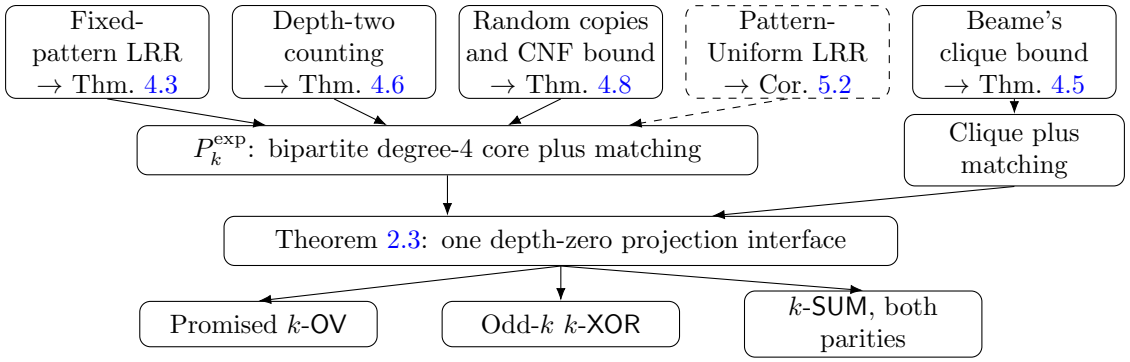


Figure 1: Each source lower-bound argument yields the indicated target result. Both pattern families have $k-1$ edges and use Theorem 2.3. The dashed arrow indicates the use of Conjecture 5.1.

copy of P. Each accepting selection identifies one present host edge for each pattern edge, with agreement at shared endpoints. OV assigns a separate array to the possible edges for each pattern edge, so its definition already enforces one choice from each array. XOR and SUM enforce these choices by equations within a single list. The direct XOR construction requires odd k; OV and SUM work for either parity. Section 3 gives the constructions and their correctness proofs.

The key property is that every bit of the constructed target instance is a constant, a source input bit or its negation. Substituting these bits into the input leaves of a target circuit produces a source circuit without adding gates or increasing depth, and with the same top-gate type. Thus each source lower bound gives a target lower bound through the same constructions (Theorem 2.3).

For a pattern with $k-1$ edges, an accepting selection uses $k-1$ objects representing edges and one fixed auxiliary object (Section 2.2). Each OV array assigned to a pattern edge accommodates N^2 possible host edges, whereas each XOR or SUM list accommodates $1+(k-1)N^2$ objects. The corresponding host sizes are $N=\Theta(\sqrt{n})$ for OV and $N=\Theta(\sqrt{n/k})$ for XOR and SUM. A source lower bound $N^a=(N^2)^{a/2}$ therefore translates into a power of n for OV or n/k for the lists, with half the source exponent before accounting for constant factors in the host size (Lemma 2.4). Our source patterns are padded bipartite expanders $P_k^{\exp}$ (Lemma 2.2) or padded cliques (Definition B.3); the target constructions remain fixed as the source lower bound changes (Figure 1).

Source lower bounds. We use the fixed-pattern LRR theorem for fixed k (Theorem 4.3). For growing k at every fixed depth, we transfer Beame's small-clique lower bound through a padded clique (Theorem 4.5). At depth two, we count minimal positive inputs for DNFs and minimal sets of edges meeting every colored copy for CNFs (Theorem 4.6). For top-disjunction depth three, the single-CNF minterm theorem bounds the probability that a fixed CNF accepts a planted core copy over a sparse random background while rejecting every proper subset of the planted edges over that same background (Theorem 4.7). A union bound over the middle CNFs yields Theorem 4.8. Pattern-Uniform LRR (Conjecture 5.1) would give the remaining growing-k bounds with an exponent linear in k.

1.2 Main results

The main target lower bounds combine the common projections with the source bounds appropriate to each regime. For reference, A–F denote Theorem 2.3, Theorem 4.3, Theorem 4.5, Theorem 4.6, Theorem 4.8 and Corollary 5.2; subscripts in constants such as ξ_D and $k_F(d)$ identify

Result	Regime / circuit class	Lower bound	Qualification
Theorem 4.3	Large fixed k, every fixed depth	$n^{\beta(k-1)}$ (k-OV); $(n/k)^{\beta(k-1)}$ (k-XOR, k-SUM)	Universal $\beta > 0$; minimum n depends on (d, k)
Theorem 4.5	Growing k, fixed depth d	$n^{\Omega_d(\min\{\sqrt{k}, \log n\})}$	Unconditional
Theorem 4.6	Growing k, depth two	$n^{\Omega(k)}$	Both orientations
Theorem 4.8	Growing k, Σ_3	n^{ck} (k-OV); $(n/k)^{ck}$ (k-XOR, k-SUM)	Every $0 < c < 1/16$; Σ_3 only
Corollary 5.2	Growing k, Π_3 or fixed $d \geq 4$	$n^{\Omega_d(k)}$	Conditional on Conjecture 5.1

Table 1: Main lower bounds for the three targets from the common subgraph framework. The growing-k rows assume $k \to \infty$ and $k = n^{o(1)}$, so $n/k = n^{1-o(1)}$. Direct k-XOR bounds require odd k, and k-OV retains its distinctness promise.

the corresponding result. Results A–E are unconditional, and F is a conditional corollary. Table 1 summarizes the resulting target bounds.

Budgets and parameter ranges. The principal budgets are $D \geq C_{\mathrm{OV}} k \log n$ for OV and $m \geq C_{\oplus} k \log(en/k)$ or $m \geq C_{\Sigma} k \log(en/k)$ for XOR or SUM, respectively, with universal construction constants (Lemma 2.4). In Theorem 4.3, the same constant β in the exponent works at every fixed depth and for every sufficiently large fixed k; the minimum required n may depend on (d, k).

For growing k, each result gives a range of k relative to n that ensures the source bound can be used at the corresponding host size N. In Theorem 4.5, we combine Beame's small-clique lower bound with the common projections for a clique padded to $k-1$ edges. The theorem holds throughout $k \to \infty$, $k \leq n^{1-\varepsilon}$ for fixed $\varepsilon > 0$; its full range uses the wider SUM budget $m \geq C_{\Sigma} k \log n$. The $k-1$ pattern edges allow only $O(\sqrt{k})$ vertices in the clique, and Beame's bound also requires at most $\log N$ clique vertices. These limits explain $\min\{\sqrt{k}, \log n\}$ in the exponent in Table 1: the displayed bound stops improving with k once $k \geq \log^2 n$.

The other source lower-bound arguments use one padded bipartite expander family. For Theorem 4.6, the requirement $N \geq k^2$ gives the target ranges $k \leq \xi_D n^{1/4}$ for OV and $k \leq \xi_D n^{1/5}$ for XOR and SUM, with universal $\xi_D > 0$. In Theorem 4.8, any fixed coefficient $0 < c < 1/16$ gives the target bounds n^{ck} for OV and $(n/k)^{ck}$ for XOR and SUM, provided k is sufficiently large and $n \geq k^{C'}$ for a constant $C' > 1$ depending on c.

Conditional and supplementary scope. Pattern-Uniform LRR (Conjecture 5.1) asks for a source exponent linear in k at a polynomial host threshold, uniformly along the expander family at each fixed depth. Corollary 5.2 then gives the remaining growing-k lower bounds with an exponent linear in k, using the same projections.

Appendix C extends the lower bounds to other dimension, row-count and bit-width budgets. The XOR lift in Appendix D.2 transfers lower bounds from odd witness sizes to even ones under explicit arithmetic and size conditions. A supplementary PARITY projection gives a tight depth-two bound for odd XOR, with further even-parameter consequences (Appendix D). The subgraph bounds are stronger for sufficiently large fixed k at depths $d \geq 3$ and for growing k in the Σ_3 range of Theorem 4.8, and also cover OV and SUM.

At the principal budget scales (within constant factors) for fixed k, with constants chosen to meet the lower-bound thresholds, elementary Σ_3 upper bounds match the lower bounds for all three

targets up to constants in the exponent, with odd k for direct XOR (Propositions E.1 and E.3). SUM uses a *block-carry verifier*, which guesses carries between blocks of bits and checks each block with a CNF. At width $m = \Theta(k \log(en/k))$ with a sufficiently large constant factor and $n \geq k^2$, its size is $(n/k)^{O(k)}$. This matches the growing-k Σ_3 lower-bound scale of Theorem 4.8 for sufficiently large k and $n \geq k^{C'}$ with a suitable constant $C' > 1$.

1.3 Related work

Direct Orthogonal Vectors lower bounds. Kane and R. Ryan Williams prove OV lower bounds for branching programs, bounded-fan-in formulas, and formulas with unbounded-fan-in symmetric gates, within $\widetilde{O}(D)$ factors of natural upper bounds in their main regimes [KW19, Thms. 2, 3, and 5]. Choudhury, Limaye, Sreenivasaiah, and Srinivasan sharpen the first two to $\Omega(n^2 D)$ for formulas and $\Omega(n^2 D/\log(nD))$ for branching programs when $D > c \log n$, for constant $c > 1$ [CLSS26, Thms. 6 and 25]. These are different computational models: a formula cannot share an intermediate result among several later gates, whereas our circuits can. Kane–Williams count wires in their symmetric-gate theorem and their worst-case depth-three conjecture [KW19, Thm. 5 and Conj. 10]. A wire count charges each connection; our size measure charges each non-input gate once, regardless of its fan-in.

Kane–Williams also give an average-case upper bound: for each fixed $p \in (0, 1)$, when every input bit is independently one with probability p, AC^0 formulas of size $n^{2-\varepsilon_p}$ for some $\varepsilon_p > 0$ are correct with probability $1 - o(1)$ [KW19, Thm. 6]. These formulas may err on exceptional inputs; our circuits must be correct on every input satisfying the promise.

The closest Σ_3 result is due to Choudhury and Sreenivasaiah. For $k \leq D$, they show that computing unpromised k-OV as a disjunction of CNFs with at most t literals per clause requires $\Omega((n/t)^k)$ CNFs feeding the top OR gate [CS26, Thm. 1.1]. For fixed k and $k \leq t = O(1)$, this gives exponent k, matching exhaustive search over witnesses. Their generalization allows bottom gates that are monotone after fixing each input's polarity, with at most t negated inputs [CS26, Thms. 1.2 and 5.12]. They also prove an unrestricted Π_3 lower bound of $2^{\Omega(n)}$ for unpromised 2-OV, at dimension $D = \Omega(n^2)$ [CS26, Thm. 1.3].

Their bound for t-CNFs also covers growing k at dimension $O(k \log n)$. At this dimension scale, Theorem 4.8 gives an $n^{\Omega(k)}$ lower bound without a restriction on clause width, in its stated range of n and k. Our OV bounds are proved on arrays with pairwise-distinct vectors; they also hold when repetitions are allowed, because a circuit correct on all arrays must be correct on the promised inputs. The same framework also gives XOR bounds for odd k and SUM bounds for both parities.

Choudhury–Limaye–Sreenivasaiah–Srinivasan give a stronger exponent for unrestricted fixed-depth circuits at polynomial dimension: for fixed $k, d \geq 2$ and $\varepsilon > 0$, their bound is $n^{k-\varepsilon}$ at $D = n^{\varepsilon}$ [CLSS26, Thm. 4 and Footnote 4]. They state this for the complement of OV, which accepts when no orthogonal tuple exists. Complementation preserves size and depth in our circuit model, so the lower bound applies to OV as well. Their reduction already produces pairwise-distinct vectors [CLSS26, Thm. 11]. Our fixed-k bound has a smaller exponent but holds at dimension $O(k \log n)$.

For monotone circuits computing the complement of 2-OV with repetitions allowed, the same authors prove an $\Omega(n^{2-\varepsilon})$ lower bound at $D = c_\varepsilon \log n$ for every fixed $0 < \varepsilon < 1$ and a suitable constant c_ε [CLSS26, Thm. 5]. Monotone circuits have no input negations; our lower bounds allow them.

Source lower bounds and broader context. Classical AC^0 lower bounds establish the limitations of constant-depth computation [Ajt83, FSS84, Yao85, Hås86]. Our source imports belong

to the subgraph-detection line of Beame, Rossman, Amano, and Li–Razborov–Rossman [Bea90, Ros08, Ama10, LRR17]. LRR's colored fixed-pattern theory and Rosenthal's average-case refinement [Ros21] relate the circuit-size exponent to the structure of the pattern. We use the fixed-pattern LRR bound for Theorem 4.3 and Beame's uniform small-clique bound for Theorem 4.5; the depth-two and Σ_3 source arguments are proved here. The conditional extension asks for an exponent linear in k once the host size is polynomial in k, with constants depending only on depth. Published fixed-pattern statements do not provide that guarantee.

Other OV–circuit connections study what faster OV algorithms would imply for circuit satisfiability and circuit lower bounds [ABDN18, RW25]. Their win–win arguments establish that at least one of two circuit lower bounds must hold, by analyzing different possibilities for OV algorithms. Our results directly lower-bound the size of circuits computing the witness-selection problems.

For XOR and SUM, generalized-birthday algorithms exploit random lists with many solutions [Wag02]. Other work studies the time needed to find a solution in random lists or in lists constructed with a hidden solution [DKK24, ASS+24]. Parity-counting hardness concerns whether the number of solutions is odd [DLSV25]. Our question is the circuit size needed to decide whether a solution exists on every input. Our proofs use no running-time assumption from this work, and worst-case circuit hardness alone does not establish hardness on the random distributions used in cryptography.

Organization. Section 2 defines the source problem, shared pattern family and projection interface; Section 3 gives all three constructions and their correctness proofs. Section 4 derives the unconditional results, ending with the minterm theorem and its depth-three consequence. Section 5 gives the conditional corollary, synthesis and open problems. Appendices A and B give details of the source bounds and pattern constructions. Appendix C extends the lower bounds to other dimension, row-count and bit-width budgets; Appendix D gives the supplementary XOR results. Appendix E proves the elementary upper bounds, and Appendix F proves the depth-three results of Section 4.4.

2 The Projection Framework and The Source Problem

We now define the source problem and formalize the common projections to the targets of Definition 1.1. Theorem 2.3 specifies how large a source host each target can accommodate and what target dimension or bit width suffices for the projection. Lemma 2.4 then translates source lower-bound exponents into target exponents. We begin with the source problem and shared pattern family, then give the circuit model and common projection construction in Section 2.2.

2.1 The source problem and shared patterns

For a finite simple loopless undirected graph P, write $v(P) := |V(P)|$, $e(P) := |E(P)|$, and $\Delta(P)$ for its maximum degree. We denote pattern edges by f and fix an arbitrary ordering $f = (f_{\mathrm{L}}, f_{\mathrm{R}})$ of each edge's endpoints, determining the order of the two host indices.

Definition 2.1 (Colored P-subgraph isomorphism)**.** Fix a pattern P and an integer $N \geq 1$. We write $P\text{-SUB}_N$ for the Boolean function deciding colored P-subgraph isomorphism. The host has vertex set $V(P) \times [N]$, with color class $\{a\} \times [N]$ for each $a \in V(P)$. Its input consists of bits

$$X_{f,u,w} \in \{0,1\}, \qquad f = (f_{\mathrm{L}}, f_{\mathrm{R}}) \in E(P), \qquad u, w \in [N]. \tag{1}$$

The variable $X_{f,u,w}$ records whether the colored host contains the edge between (f_{L}, u) and (f_{R}, w). The function accepts exactly when there is a map $\phi : V(P) \to [N]$ such that

$$X_{f,\phi(f_{\mathrm{L}}),\phi(f_{\mathrm{R}})} = 1 \qquad \text{for every } f \in E(P). \tag{2}$$

An accepting map ϕ specifies a *colored copy* of P: its vertices are $(a, \phi(a))$ for $a \in V(P)$, and its edges are those required by (2).

The projections will describe colored copies by their edges. Choose one present host edge per pattern edge. For each non-isolated pattern vertex a, require all chosen edges incident to a to use the same host vertex (a, u). This common index defines $\phi(a) = u$. Any isolated pattern vertices may be assigned arbitrary indices in $[N]$.

The map ϕ assigns indices within color classes. For example, if $a \neq b$ and $\phi(a) = \phi(b) = 1$, the selected host vertices are $(a, 1)$ and $(b, 1)$, which are distinct because their colors differ. Thus ϕ need not be injective into $[N]$, although the map $a \mapsto (a, \phi(a))$ into the host is always injective.

This is exactly the structured colored function of Li, Razborov, and Rossman [LRR17, Sec. 5, p. 960]; Proposition A.1 gives the coordinate bijection. Their host parameter n is our N. Our n denotes the number of vectors in each OV array or the number of entries in an XOR or SUM list.

The shared padded family. The common projections (Theorem 2.3) require a source pattern with exactly $k-1$ edges. The shared family below starts with a bipartite 4-regular expander core and adds at most three disjoint edges to reach this count. For a core graph F, let J_t be a matching of $t \geq 0$ edges on new vertices, with J_0 empty. We write $P = F \mathbin{\dot\cup} J_t$ for their disjoint union.

Following the endpoint convention in Section 2.1, choose an order $(f_{\mathrm{L}}, f_{\mathrm{R}})$ for each edge f. Keep the orders chosen for F when forming P, and choose those for the matching edges arbitrarily. Fix once and for all a universal threshold $k_0 \geq 600$ large enough for the following construction to work for every $k \geq k_0$; Appendix B.1 shows that such a finite threshold exists.

Lemma 2.2 (Padded bipartite expander patterns)**.** *For every integer $k \geq k_0$, there is a simple loopless graph*

$$P_k^{\mathrm{exp}} = F_k^{\mathrm{exp}} \mathbin{\dot\cup} J_{t_k} \tag{3}$$

with exactly $k-1$ edges, no isolated vertices, and maximum degree 4, where F_k^{exp} is a simple connected bipartite 4-regular expander and $0 \leq t_k < 4$. Fixing either bipartition side as I_k, one has

$$e(F_k^{\mathrm{exp}}) > k - 5, \qquad |I_k| = \left\lfloor \frac{k-1}{4} \right\rfloor > \frac{k-1}{4} - 1. \tag{4}$$

The construction and proof are in Appendix B.1. Fix one such family for the rest of the paper. It supports fixed-pattern hardness in Section 4.1.1, elementary depth-two counting in Theorem 4.6, the independent-side argument in Theorem 4.8, and the conjecture behind Corollary 5.2. The proofs of Theorems 4.6 and 4.8 use the displayed structure without using expansion. For growing k at arbitrary fixed depth, Theorem 4.5 instead uses a clique padded by a matching; its degree may grow. These are the two family branches of Figure 1.

2.2 The circuit model and common projection construction

Circuit conventions. All logarithms are base two. Circuits have unbounded-fan-in AND/OR gates and negations only at input leaves; $\mathrm{size}_d(g)$ is the minimum number of non-input gates in

such a circuit of depth at most d computing g. For the depth-three orientations Σ_3 and Π_3 defined in the Introduction, flattening adjacent gates of the same type gives

$$\mathrm{size}_3(g) = \min\{\mathrm{size}_{\Sigma_3}(g), \mathrm{size}_{\Pi_3}(g)\}. \tag{5}$$

For a CNF T, let $s(T)$ be its clause count. Write M for the number of input variables; for the source function P-SUB$_N$, we denote this count by $M_{\mathrm{src}} := e(P)N^2$. A Σ_3 circuit of size S on M variables can be written as $\bigvee_{i=1}^{t} T_i$ with

$$t \le S + 2M, \qquad s(T_i) \le S + 2M \quad (i \in [t]). \tag{6}$$

Indeed, each input to the top OR or to a middle AND is a gate output or a literal. Repeated inputs to the same gate can be deleted, whether they are copies of one gate output or of one literal. At most S gate outputs and $2M$ distinct literals remain available, giving both bounds. The dual statement holds for Π_3.

Target conventions. We use the targets defined in Definition 1.1. For the promised k-OV function, size_d denotes the minimum size of a circuit that is correct on every promised input.

The OV promise makes each array a set in an arbitrary order. A circuit correct on unrestricted arrays is also correct on promised arrays, so our lower bounds apply to the unrestricted variant. List indices, by contrast, may carry equal values: XOR and SUM may repeat padding values, while OV must keep the vectors within each array pairwise distinct.

Depth-zero projections. For possibly promised Boolean functions F and G, write $F \le_{\mathrm{proj}} G$ if a map R sends the promise domain of F into that of G, satisfies $F(x) = G(R(x))$ on every promised input, and has each output bit equal to a constant, an input bit, or its negation.

Projection monotonicity. If $F \le_{\mathrm{proj}} G$, then $\mathrm{size}_d(G) \ge \mathrm{size}_d(F)$ for every depth d, also within a fixed orientation such as Σ_3 or Π_3. Indeed, substituting the projection's constants and literals at the input leaves of a circuit for G gives a circuit for F: the promise is respected, and propagating constants cannot increase size or depth or change the required orientation.

Selecting edges. Fix an integer $k \ge 2$ and a simple loopless pattern P with $e(P) = k - 1$, using the endpoint orders $f = (f_{\mathrm{L}}, f_{\mathrm{R}})$ fixed in Section 2.1.

By Definition 2.1, choosing one present host edge for each pattern edge, with agreeing labels at shared endpoints, is equivalent to choosing a colored copy of P. Our target input has an *edge object* representing each possible host edge. The objects for one fixed pattern edge form an *edge group*. A target witness must select one edge object from each of the $k-1$ groups and one *fixed auxiliary object*, making k selections altogether. Additional *padding objects* fill the input to its prescribed length n and must be excluded from accepting witnesses.

For $f = (f_{\mathrm{L}}, f_{\mathrm{R}}) \in E(P)$ and $(u, w) \in [N]^2$, we use an edge object $c_{f,u,w}$ (corresponding to source bit $X_{f,u,w}$) to represent the possible host edge from (f_{L}, u) to (f_{R}, w). We will later show that each construction uses the absence bit $\overline{X}_{f,u,w} := 1 - X_{f,u,w}$ to check edge presence, encoding this object as a vector for OV and XOR or an integer for SUM. Write α for the fixed auxiliary object. Before padding, the objects are

$$\{\alpha\} \,\dot\cup \bigcup_{f \in E(P)} \{c_{f,u,w} : (u, w) \in [N]^2\}. \tag{7}$$

For k-XOR and k-SUM, these objects occupy one indexed list and number $1 + e(P)N^2$; under $1 + e(P)N^2 \le n$, padding entries fill the list to length n. For k-OV, the auxiliary object and the $e(P)$ edge groups become k separate arrays: the auxiliary array contains one vector before padding, and each edge array contains N^2 vectors before padding. Under $N^2 \le n$, padding vectors fill every array to length n.

For odd-k XOR and for SUM, the fixed auxiliary object enters equations that, together with the required k selections and the padding checks, force exactly one edge object per group (Sections 3.2 and 3.3). OV already requires one vector per array, so the auxiliary object has no role in enforcing one edge choice per group. Its array lets the same $(k-1)$-edge pattern give k selections for all three targets.

For each pattern vertex a, define lab_a on triples (f, u, w) with $f \in E(P)$ incident to a and $(u, w) \in [N]^2$. It returns the host label at endpoint a of the edge represented by $c_{f,u,w}$:

$$\mathrm{lab}_a(f, u, w) := \begin{cases} u, & a = f_{\mathrm{L}}, \\ w, & a = f_{\mathrm{R}}. \end{cases} \tag{8}$$

Put $\ell_N := \lceil \log N \rceil$ and let $\mathrm{bit}_t(x)$ be the t-th bit of the ℓ_N-bit encoding of $x - 1$, for $x \in [N]$ and $t \in [\ell_N]$.

Selected edges sharing a pattern vertex must agree on its host label to define a single source map ϕ. It suffices to compare every incident edge with one reference edge, since equality is transitive. Accordingly, fix one reference incident edge $\rho(a)$ for each non-isolated vertex a, and let

$$\mathcal{I}(P) := \{(a, g) : a \in V(P),\ \deg_P(a) \ge 1,\ g \ni a,\ g \ne \rho(a)\}. \tag{9}$$

Thus $\mathcal{I}(P)$ indexes one comparison for every non-reference incidence, and

$$|\mathcal{I}(P)| = \sum_{a:\deg_P(a)\ge 1} (\deg_P(a) - 1) \le \sum_{a \in V(P)} \deg_P(a) = 2e(P). \tag{10}$$

Degree-one vertices contribute no non-reference incidence and zero to this sum.

Correctness requirements. To make target acceptance equivalent to the existence of a colored copy of P, the constructions must enforce the following requirements.

- *Selection.* The witness must contain the auxiliary object and exactly one edge object per group, so the k choices cover every pattern edge once. Otherwise, several choices could come from one group while another pattern edge goes unchecked.
- *Edge presence.* Every selected $c_{f,u,w}$ must satisfy $X_{f,u,w} = 1$, since edge objects represent all possible host edges, including absent ones.
- *Endpoint agreement.* The selected edges must agree on the host label at each shared pattern vertex, so that they fit together into a single source map (Definition 2.1).
- *Padding exclusion.* The objects added to fill the input must not create accepting witnesses unrelated to a source copy. For example, k zero padding entries would make an XOR or SUM list accept even if the source host had no edges.
- *Vector distinctness (OV).* We ensure pairwise-distinct vectors within each array without changing which tuples are orthogonal.

The constructions in Section 3 show how each target enforces these conditions.

2.3 Master projections and host substitutions

We now quantify how a circuit lower bound for colored subgraph isomorphism yields lower bounds for the three targets. A larger source host strengthens the resulting bound, so we want to choose N as large as the target's arrays or list can accommodate. Theorem 2.3 states the capacity and encoding-budget requirements; Lemma 2.4 then chooses the hosts and converts a source bound N^a into a power of n for OV or n/k for XOR and SUM.

Notation for bounds. We generally write source lower bounds with exponent rate γ and target lower bounds with rate β or c; Lemma 2.4 explains how these rates are related. The universal constants $C_{\mathrm{OV}}, C_{\oplus}, C_{\Sigma}$ in Theorem 2.3 specify the dimension, row-count, and bit-width requirements for the projections.

The next theorem uses the native budgets $D_0(P,N,n)$, $m_{\oplus}(P,N)$ and $m_{\Sigma}(P,N)$ defined in Equations (21), (27) and (31), respectively.

Theorem 2.3 (Master projections)**.** *There are universal constants $C_{\mathrm{OV}}, C_{\oplus}, C_{\Sigma} > 0$ such that the following holds. Let $k \geq 2$, let P be a simple loopless pattern with $e(P) = k-1$, and fix an endpoint order for each edge. Then the following depth-zero projections exist.*

(i) If $N^2 \leq n$ and $D \geq D_0(P,N,n)$, then P-SUB$_N$ $\leq_{\mathrm{proj}}$ k-OV$_{n,D,k}$, where $D_0(P,N,n) \leq C_{\mathrm{OV}} k \log(2n)$. Moreover, $D_0(P,N,n) \geq \lceil \log n \rceil$, so the hypothesis on D automatically ensures $n \leq 2^D$.

(ii) If k is odd, $1 + e(P)N^2 \leq n$, and $m \geq m_{\oplus}(P,N)$, then P-SUB$_N$ $\leq_{\mathrm{proj}}$ k-XOR$_{n,m,k}$, where $m_{\oplus}(P,N) \leq C_{\oplus} k \log(2N)$.

(iii) If $1 + e(P)N^2 \leq n$ and $m \geq m_{\Sigma}(P,N)$, then P-SUB$_N$ $\leq_{\mathrm{proj}}$ k-SUM$_{n,m,k}$, where $m_{\Sigma}(P,N) \leq C_{\Sigma} k \log(2kN)$. This projection has no parity restriction.

Sections 3.1 to 3.3 prove the respective projections and budget bounds. We choose $C_{\mathrm{OV}}, C_{\oplus}, C_{\Sigma}$ once, with $C_{\oplus} \geq 1$, to satisfy both Theorem 2.3 and the budget estimates in Lemma 2.4. The host N remains free, preserving OV's larger capacity.

Lemma 2.4 (Host substitutions and lower-bound transfer)**.** *Let $k \geq 2$ and $n \geq 16k$, and define the k-OV host size and the common host size for k-XOR (with odd k) and k-SUM by*

$$N_{\mathrm{OV}} := \lfloor \sqrt{n} \rfloor, \qquad N_{\oplus,\Sigma} := \left\lfloor \sqrt{\frac{n-1}{k-1}} \right\rfloor. \tag{11}$$

For every simple loopless pattern P with $e(P) = k-1$, these choices satisfy

$$N_{\mathrm{OV}}^2 \leq n, \qquad 1 + (k-1)N_{\oplus,\Sigma}^2 \leq n, \tag{12}$$

and

$$N_{\mathrm{OV}} \geq \frac{1}{2}\sqrt{n}, \qquad \log N_{\mathrm{OV}} \geq \frac{1}{4}\log n, \tag{13}$$

$$N_{\oplus,\Sigma} \geq \frac{1}{2}\sqrt{\frac{n}{k}}, \qquad \log N_{\oplus,\Sigma} \geq \frac{1}{4}\log \frac{n}{k}. \tag{14}$$

At these host sizes, the budgets have the following bounds in n and k:

$$D_0(P, N_{\mathrm{OV}}, n) \leq C_{\mathrm{OV}} k \log n, \tag{15}$$

$$m_{\oplus}(P, N_{\oplus,\Sigma}) \leq C_{\oplus} k \log(en/k), \tag{16}$$

$$m_{\Sigma}(P, N_{\oplus,\Sigma}) \leq C_{\Sigma} k \log n. \tag{17}$$

If in addition $n \geq k^3$, then the last inequality sharpens to

$$m_\Sigma(P, N_{\oplus,\Sigma}) \leq C_\Sigma k \log(en/k). \tag{18}$$

For $Q \in \{k\text{-OV}_{n,D,k}, k\text{-XOR}_{n,m,k}, k\text{-SUM}_{n,m,k}\}$, let N_Q be its host in Equation (11) and put $B_Q := n$ for k-OV and $B_Q := n/k$ for k-XOR or k-SUM. At the budgets above, with odd k for k-XOR, the following holds for every fixed depth d: if $\mathrm{size}_d(P\text{-SUB}_{N_Q}) \geq N_Q^a$ for some $a > 0$, then

$$\mathrm{size}_d(Q) \geq (B_Q/4)^{a/2} \geq B_Q^{a/4}. \tag{19}$$

To retain a target rate close to half the source rate, suppose $a = \gamma s$ with $\gamma, s > 0$. Then $\mathrm{size}_d(Q) \geq B_Q^{cs}$ for any $0 < c < \gamma/2$ whenever

$$B_Q^{\gamma/2-c} \geq 2^\gamma. \tag{20}$$

Both transfer statements also hold separately with size_{Σ_3} or size_{Π_3} in place of size_d.

Proof. Rounding down the square roots in Equation (11) gives the inequalities in Equation (12): the edge objects fit in the OV arrays, and the auxiliary object and all edge objects fit in the XOR or SUM list.

For the lower estimates, $n \geq 16k$ and $k \geq 2$ give $(n-1)/(k-1) \geq n/k \geq 16$. Applying $\lfloor x \rfloor \geq x/2$ for $x \geq 2$ to the square roots in Equation (11) yields the lower bounds for N_{OV} and $N_{\oplus,\Sigma}$ in Equations (13) and (14). Since B_Q equals n or n/k, we have $B_Q \geq 16$ and hence $\log N_Q \geq \frac{1}{2} \log B_Q - 1 \geq \frac{1}{4} \log B_Q$.

To express the budgets in n and k, use $\log(2n) \leq 2 \log n$ and $N_{\oplus,\Sigma} \leq \sqrt{2n/k}$. Under $n \geq 16k$, the latter gives $\log(2N_{\oplus,\Sigma}) \leq \log(en/k)$ and $\log(2kN_{\oplus,\Sigma}) \leq \log n$. If $n \geq k^3$, then $kN_{\oplus,\Sigma} \leq \sqrt{2kn} \leq \sqrt{2}\, n/k$, so $\log(2kN_{\oplus,\Sigma}) \leq 2 \log(en/k)$. These conversions change the coefficients in the budget bounds of Theorem 2.3 by at most a factor of two. Choosing the construction constants large enough to cover the estimates both before and after substitution proves Equations (15) to (18), with the last bound requiring $n \geq k^3$.

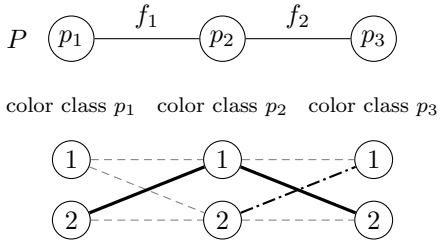
Apply Theorem 2.3 and projection monotonicity to obtain $\mathrm{size}_d(Q) \geq N_Q^a$. By Equations (13) and (14), we have $N_Q \geq \sqrt{B_Q}/2$; together with $B_Q \geq 16$, this proves Equation (19). For $a = \gamma s$, the first target bound in Equation (19) equals $(B_Q^{\gamma/2}/2^\gamma)^s$, which is at least B_Q^{cs} under Equation (20). The same proof applies separately to Σ_3 and Π_3, since the projections preserve orientation. □

The lemma supplies the capacity, budget and exponent conversion. Each application must still check its source pattern, the minimum required host size and any upper restriction on the host size. We call the target thresholds obtained from Equations (15), (16) and (18) the *principal budgets*. The sharper SUM bound requires $n \geq k^3$; the general SUM bound (17) is needed for the full range of Theorem 4.5.

3 Three Constructions

We prove the three parts of Theorem 2.3 using the objects in (7). Each accepting witness must select the auxiliary object and one present edge object per group, with agreeing endpoint labels and no padding. These conditions give a colored copy of P by Definition 2.1. We use the endpoint labels lab_a and reference comparisons $\mathcal{I}(P)$ from Equations (8) and (9). OV also requires distinct vectors within each array.

Running example. We use a single source instance to illustrate the three constructions and their correctness checks. Figure 2 shows the source, and Figures 3 to 5 show its OV, XOR and SUM encodings, respectively. Set $k = 3$, $N = 2$, and let P be the path with edges $f_1 = (p_1, p_2)$, $f_2 = (p_2, p_3)$. For the reference comparisons in (9), choose $\rho(p_2) = f_1$. Thus $\mathcal{I}(P) = \{(p_2, f_2)\}$ and $\ell_N = 1$: the comparison tests the second label of an edge object for f_1 against the first label of an edge object for f_2. Write $X^i_{uw} := X_{f_i,u,w}$. The unique copy selects $X^1_{21} = X^2_{12} = 1$; the present edge $X^2_{21} = 1$ gives a different label at the shared endpoint. OV uses array size $n = 4$; XOR and SUM use list length $n = 12$, leaving three padding indices. Repeated padding values are legal in XOR and SUM.



X^i_{uw}	11	12	21	22
$i = 1$	0	0	**1**	0
$i = 2$	0	**1**	1	0

Example with a unique copy: $(2, 1, 2)$.

Figure 2: The source instance used in all three constructions, with $N = 2$. Heavy edges form the unique copy $(2, 1, 2)$; the dash-dotted edge is present but disagrees at the shared endpoint, and gray dashed edges are absent. The table gives all eight source bits.

3.1 k-Orthogonal Vectors

We encode the objects in (7) as binary vectors: the auxiliary object α is represented by v_α in an array U_α, and each edge object $c_{f,u,w}$ by $v_{f,u,w}$ in an array U_f, one array for each $f \in E(P)$. Fill every array to size n with padding vectors. Once padding is excluded, choosing one vector per array gives the auxiliary vector and one edge vector per group.

A choice of one vector per array is orthogonal exactly when, in each coordinate, the product of the k selected entries is zero (Definition 1.1). We therefore design coordinates of these vectors so that an absent edge, endpoint disagreement or padding choice makes some coordinate equal to 1 on all selected vectors. Additional identifier coordinates ensure distinctness without changing orthogonality. In Figure 3, columns are vectors grouped by array and rows are coordinates; the starred columns form an orthogonal witness.

Edge presence and padding exclusion. We introduce one coordinate q_α to force the choice of v_α from U_α. It is 0 on v_α and 1 on every other vector. Selecting v_α makes the product on q_α zero. Selecting a padding vector from U_α instead makes all k selected entries on q_α equal to 1, so those vectors cannot form an orthogonal witness.

For each pattern edge $f \in E(P)$, introduce one coordinate q_f to check that the choice from U_f represents a present edge, giving $e(P) = k - 1$ such coordinates. Set q_f to $\overline{X}_{f,u,w}$ on $v_{f,u,w}$ and to 1 everywhere else. If $v_{f,u,w}$ is selected, the other $k - 1$ selected vectors contribute 1, so the product on q_f is $\overline{X}_{f,u,w}$: a present edge passes with product 0, and an absent edge is rejected with product 1. Choosing padding from U_f also gives product 1, excluding padding from that array.

The second and fourth selections in the diagnostic table of Figure 3 illustrate these rejections. In the second, v^{11}_1 represents an absent edge, so the product on q_{f_1} is 1; in the fourth, δ^1 is padding in U_α, so the product on q_α is 1.

Role	coordinate	U_α				U_{f_1}				U_{f_2}			
		$\boldsymbol{v_\alpha^\star}$	δ^1	δ^2	δ^3	v_1^{11}	v_1^{12}	$\boldsymbol{v_1^{21\star}}$	v_1^{22}	v_2^{11}	$\boldsymbol{v_2^{12\star}}$	v_2^{21}	v_2^{22}
Padding	q_α	0	1	1	1	1	1	1	1	1	1	1	1
Edges	q_{f_1}	1	1	1	1	$\boxed{\overline{X}^1_{11(1)}}$	$\boxed{\overline{X}^1_{12(1)}}$	$\boxed{\overline{X}^1_{21(0)}}$	$\boxed{\overline{X}^1_{22(1)}}$	1	1	1	1
Edges	q_{f_2}	1	1	1	1	1	1	1	1	$\boxed{\overline{X}^2_{11(1)}}$	$\boxed{\overline{X}^2_{12(0)}}$	$\boxed{\overline{X}^2_{21(0)}}$	$\boxed{\overline{X}^2_{22(1)}}$
Agreement	C^{10}	1	1	1	1	0	1	0	1	1	1	0	0
Agreement	C^{01}	1	1	1	1	1	0	1	0	0	0	1	1
Distinctness	id_1^E	0	0	0	0	0	0	1	1	0	0	1	1
Distinctness	id_2^E	0	0	0	0	0	1	0	1	0	1	0	1
Distinctness	id_1^A	0	0	1	1	0	0	0	0	0	0	0	0
Distinctness	id_2^A	0	1	0	1	0	0	0	0	0	0	0	0

Selected vectors	diagnostic	outcome
$\{v_\alpha, v_1^{21}, v_2^{12}\}$	every coordinate product is 0	accept
$\{v_\alpha, v_1^{11}, v_2^{12}\}$	product 1 on q_{f_1}	absent-edge rejection
$\{v_\alpha, v_1^{21}, v_2^{21}\}$	C^{10}-product 0, C^{01}-product 1	endpoint disagreement
$\{\delta^1, v_1^{21}, v_2^{12}\}$	product 1 on q_α	padding rejection

Figure 3: The complete 9×12 toy k-OV matrix, where $v_i^{uw} := v_{f_i,u,w}$. Rows C^{10}, C^{01} compare the endpoint labels proposed by f_1 and f_2 at p_2, using f_1 as the reference edge. They abbreviate $C^{10}_{p_2,f_2,1}, C^{01}_{p_2,f_2,1}$; the final index 1 denotes the only label bit at $N = 2$. Stars mark the orthogonal witness. Gray entries are fixed nonzero constants; boxes display absence literals $\overline{X}^i_{uw}$ from Figure 2, with their evaluated bits in parentheses. The four rows labeled "Distinctness" distinguish positions without affecting orthogonality.

Endpoint agreement. To reject endpoint disagreements, introduce two coordinates $C^{10}_{a,g,t}$ and $C^{01}_{a,g,t}$ for each reference comparison $(a, g) \in \mathcal{I}(P)$ and bit position $t \in [\ell_N]$, where $\ell_N := \lceil \log N \rceil$ as in Section 2.2. This gives $2|\mathcal{I}(P)|\ell_N$ comparison coordinates. For a fixed (a, g), put $f = \rho(a)$. Coordinate $C^{10}_{a,g,t}$ assigns $\mathrm{bit}_t(\mathrm{lab}_a(f, u, w))$ to edge vectors in U_f and $1 - \mathrm{bit}_t(\mathrm{lab}_a(g, u, w))$ to edge vectors in U_g; $C^{01}_{a,g,t}$ complements these assignments on edge vectors. Padding and all other entries are assigned 1.

When the k selected vectors contain no padding, write x, y for the two corresponding label bits proposed by the selected vectors. The products on $C^{10}_{a,g,t}$ and $C^{01}_{a,g,t}$ are $x(1 - y)$ and $(1 - x)y$, so both vanish exactly when $x = y$. Thus the reference comparisons enforce agreement of full endpoint labels.

In the source example of Figure 2, f_1 and f_2 share p_2, so a colored copy must use the same host vertex of color p_2 on both edges. We therefore compare the second index of v_1^{uw} with the first index of v_2^{uw}. At $N = 2$, labels $1, 2$ are encoded as bits $0, 1$, respectively, since we encode each label minus one.

Consider the third selection $\{v_\alpha, v_1^{21}, v_2^{21}\}$ in Figure 3. Both represented edges are present, but v_1^{21} proposes label 1 at p_2 while v_2^{21} proposes label 2, giving $x = 0, y = 1$. In row C^{10}, these edge columns contain $x = 0$ and $1 - y = 0$; in row C^{01}, they contain $1 - x = 1$ and $y = 1$. The auxiliary vector has entry 1 on both rows, so the coordinate products are $1 \cdot 0 \cdot 0 = 0$ and $1 \cdot 1 \cdot 1 = 1$, respectively. Thus C^{01} rejects this selection because the edges disagree at p_2.

For the pair v_1^{12}, v_2^{12}, the proposed labels at p_2 are $2, 1$, giving $x = 1, y = 0$. These edge columns contain $(1, 1)$ on C^{10} and $(0, 0)$ on C^{01}. Selecting them with v_α therefore gives products 1 and

0, respectively, so C^{10} rejects the disagreement. This selection also fails the edge check q_{f_1}, since $X^1_{12} = 0$ in the source example.

Vector distinctness. The edge-check and comparison coordinates may give equal vectors at different positions of an array. In Figure 3, columns v_1^{12}, v_1^{22} agree on all these rows, but id_1^E distinguishes them. We append position identifiers in two blocks of $\ell_n := \lceil \log n \rceil$ coordinates, giving $2\ell_n$ additional coordinates to meet the distinctness promise in Definition 1.1.

Let $\mathrm{code}(j)$ encode $j-1$ in ℓ_n bits. Append $(\mathrm{code}(j), 0^{\ell_n})$ to the j-th vector in each edge array and $(0^{\ell_n}, \mathrm{code}(j))$ to the j-th vector in U_α, including padding vectors. Distinct indices have distinct codes, so these coordinates distinguish all n vectors within each array for every source input.

These identifiers preserve orthogonality: every coordinate in the first block is zero throughout U_α, and every coordinate in the second is zero throughout each edge array. Since $k \geq 2$, there is at least one edge array. Thus every choice of one vector per array has product zero in each added coordinate. When $n = 1$, the blocks are empty and each array already contains just one vector.

Dimension. There are $1 + e(P) = k$ coordinates for edge presence and padding exclusion, $2|\mathcal{I}(P)|\ell_N$ for endpoint agreement and $2\ell_n$ for identifiers. Hence the vectors have dimension

$$D_0(P, N, n) := k + 2|\mathcal{I}(P)|\ell_N + 2\ell_n. \tag{21}$$

For the running example in Figure 3, this gives $D_0(P, 2, 4) = 3 + 2 + 4 = 9$.

Proof of Theorem 2.3(i). Suppose first that the source instance X is accepting, and let $\phi : V(P) \to [N]$ be an accepting map as in Definition 2.1. For each pattern edge $f = (f_\mathrm{L}, f_\mathrm{R})$, select $v_{f,\phi(f_\mathrm{L}),\phi(f_\mathrm{R})}$ from U_f; also select v_α from U_α. The selected edge vectors represent exactly the edges of the colored copy specified by ϕ. The resulting selection has product zero on q_α and every q_f, since no padding is chosen and every represented edge is present. At each shared endpoint a, all selected incident edges have label $\phi(a)$, so the comparison products are zero as well. The identifier coordinates also have zero product. Thus all coordinate products vanish, and the OV instance accepts.

Conversely, suppose the constructed OV instance accepts, and choose one vector from each array whose coordinate products are all zero. Coordinate q_α forces the choice of v_α, and each q_f forces the choice of an edge vector v_{f,u_f,w_f} with $X_{f,u_f,w_f} = 1$. The comparison coordinates ensure that all labels proposed at a non-isolated vertex a agree with the label on its reference edge $\rho(a)$. Define $\phi(a)$ to be this common label, assigning an arbitrary label in $[N]$ to each isolated vertex. Then $\phi(f_\mathrm{L}) = u_f$ and $\phi(f_\mathrm{R}) = w_f$ for every f. Since every represented edge is present, ϕ satisfies (2), and the source instance accepts.

The capacity $N^2 \leq n$ ensures that each edge array, and also the auxiliary array, can be filled to size n. The identifier coordinates make the vectors within every array distinct for every source input. Moreover, $D_0(P, N, n) \geq 2\ell_n \geq \ell_n$ gives $n \leq 2^{D_0(P,N,n)}$, so the construction satisfies the OV promise at its native dimension.

Only edge-check entries depend on the source input, each as $\overline{X}_{f,u,w}$; all other entries are constants. Hence the construction is a depth-zero projection. Appending zero coordinates preserves correctness and distinctness for every $D \geq D_0(P, N, n)$. Finally, Equation (10) gives

$$D_0(P, N, n) = O(k \log(2N) + \log(2n)) = O(k \log(2n)), \tag{22}$$

using $N^2 \leq n$ in the last step. □

3.2 Odd k-XOR

We encode the auxiliary object α and all $e(P)N^2$ edge objects $c_{f,u,w}$ in (7) as binary vectors in a single indexed list, then add padding vectors to reach length n. The target accepts exactly when k distinct list indices have vector XOR zero (Definition 1.1); equal vectors at different indices are allowed. Unlike OV, choosing k indices does not itself enforce one edge choice per edge group. We use parity equations, together with odd k, to force the auxiliary object and exactly one edge object per edge group, then check edge presence and endpoint agreement.

The vectors form the columns of a matrix whose rows are the equations below; all unspecified entries are zero. In Figure 4, columns are labeled by their objects, and stars mark the accepting selection. Write $S \subseteq [n]$ for a set of k selected indices.

Role	row	$\boldsymbol{\alpha^\star}$	c_1^{11}	c_1^{12}	$\boldsymbol{c_1^{21\star}}$	c_1^{22}	c_2^{11}	$\boldsymbol{c_2^{12\star}}$	c_2^{21}	c_2^{22}	δ_1	δ_2	δ_3
Selection	S_{f_1}	1	1	1	1	1	0	0	0	0	0	0	0
Selection	S_{f_2}	1	0	0	0	0	1	1	1	1	0	0	0
Edges	H_{f_1}	0	$\overline{X}^1_{11(1)}$	$\overline{X}^1_{12(1)}$	$\overline{X}^1_{21(0)}$	$\overline{X}^1_{22(1)}$	0	0	0	0	0	0	0
Edges	H_{f_2}	0	0	0	0	0	$\overline{X}^2_{11(1)}$	$\overline{X}^2_{12(0)}$	$\overline{X}^2_{21(0)}$	$\overline{X}^2_{22(1)}$	0	0	0
Agreement	$C_{p_2,f_2,1}$	0	0	1	0	1	0	0	1	1	0	0	0
Padding	A_{pad}	0	0	0	0	0	0	0	0	0	1	1	1

Selected objects	unique nonzero row XOR	outcome
$\{\alpha, c_1^{21}, c_2^{12}\}$	none	accept
$\{\alpha, c_1^{11}, c_2^{12}\}$	$H_{f_1} = 1$	one absent edge
$\{\alpha, c_1^{21}, c_2^{21}\}$	$C_{p_2,f_2,1} = 1$	endpoint disagreement
$\{\delta_1, \delta_2, \delta_3\}$	$A_{\text{pad}} = 1$	odd padding count

Figure 4: The complete 6×12 toy k-XOR matrix, where $c_i^{uw} := c_{f_i,u,w}$ and $\delta_1, \delta_2, \delta_3$ label the padding indices. Row $C_{p_2,f_2,1}$ compares the label bits proposed at p_2 by $f_1 = \rho(p_2)$ and f_2; its final index denotes the only label bit at $N = 2$. Stars mark the accepting selection. Gray entries are fixed nonzero constants; boxes show absence literals from Figure 2, with their evaluated bits in parentheses.

Selection and padding exclusion. Introduce one selection row S_f for each $f \in E(P)$, giving $e(P) = k - 1$ rows. It is 1 on the auxiliary object and every edge object in group f. Selecting the auxiliary object and one edge object per group gives two ones in each selection row, hence XOR zero. Add one padding row A_{pad}, equal to 1 on every padding vector, so that an odd number of padding choices is rejected.

To see why these rows force the intended selection when k is odd, let $\zeta \in \{0, 1\}$ indicate whether S contains the auxiliary object, let c_f count its edge objects in group f, and let d count its padding objects. Zero XOR on the selection and padding rows gives

$$c_f \equiv \zeta \pmod 2 \quad (f \in E(P)), \qquad d \equiv 0 \pmod 2, \tag{23}$$

while

$$\zeta + \sum_f c_f + d = k = e(P) + 1. \tag{24}$$

If $\zeta = 0$, every c_f and d is even, so the selected set has even size, contradicting odd k. Thus $\zeta = 1$, and every c_f is a positive odd integer. The auxiliary object and at least one object from each of the $k-1$ groups already account for all k indices, forcing $c_f = 1$ for every f and $d = 0$.

In Figure 4, the accepting selection has two ones per selection row; choosing all three padding indices instead gives XOR $1 \oplus 1 \oplus 1 = 1$ on A_{pad}.

Edge presence. Introduce one edge-check row H_f for each $f \in E(P)$, giving $e(P)$ rows. Set it to $\overline{X}_{f,u,w}$ on edge object $c_{f,u,w}$ in group f. Once the selection and padding rows have XOR zero, exactly one edge object c_{f,u_f,w_f} is selected from this group, so

$$\bigoplus_{j \in S} H_f(j) = 1 - X_{f,u_f,w_f}. \tag{25}$$

Thus this row has XOR zero for a present edge and XOR one for an absent edge. The second selection of Figure 4 fails on H_{f_1} because $X^1_{11} = 0$.

Endpoint agreement. For each reference comparison $(a, g) \in \mathcal{I}(P)$ and bit position $t \in [\ell_N]$, introduce a row $C_{a,g,t}$, giving $|\mathcal{I}(P)|\ell_N$ rows. Put $f = \rho(a)$. On an edge object $c_{h,u,w}$ in either group $h \in \{f, g\}$, this row is $\text{bit}_t(\text{lab}_a(h, u, w))$, the corresponding bit of the proposed endpoint label. Recall that $\text{bit}_t(x)$ is the t-th bit of the ℓ_N-bit encoding of $x - 1$. With exactly one selected edge object per group, the row XOR is

$$\text{bit}_t(\text{lab}_a(f, u_f, w_f)) \oplus \text{bit}_t(\text{lab}_a(g, u_g, w_g)), \tag{26}$$

which vanishes exactly when those two bits agree. Testing all bits makes the two labels equal, and comparing every incident edge with its reference edge makes all labels at a shared endpoint agree.

In the third selection of Figure 4, c^{21}_1 proposes label 1 at p_2 and c^{21}_2 proposes label 2. Since these labels are encoded as bits $0, 1$, the row $C_{p_2,f_2,1}$ has XOR $0 \oplus 1 = 1$ and rejects their disagreement. In the accepting selection, c^{21}_1 and c^{12}_2 both propose label 1, giving XOR $0 \oplus 0 = 0$ on that row.

Row count. There are $e(P)$ selection rows, $e(P)$ edge-check rows, $|\mathcal{I}(P)|\ell_N$ comparison rows and one padding row. Hence

$$m_{\oplus}(P, N) := 2e(P) + |\mathcal{I}(P)|\ell_N + 1 = O(k \log(2N)), \tag{27}$$

using Equation (10).

Proof of Theorem 2.3(ii). Let $\phi : V(P) \to [N]$ be an accepting source map as in Definition 2.1. Select the indices of α and $c_{f,\phi(f_{\text{L}}),\phi(f_{\text{R}})}$ for each $f \in E(P)$. These k distinct indices contain no padding, and every selection row has two ones. Edge presence and endpoint agreement give XOR zero on the remaining rows, so the constructed XOR instance accepts.

Conversely, suppose k distinct indices have vector XOR zero. The selection and padding equations force them to represent α and exactly one edge object c_{f,u_f,w_f} per group, with no padding. Each row H_f forces $X_{f,u_f,w_f} = 1$, and the comparison rows make all proposed labels at a non-isolated vertex agree with the label on its reference edge. Define $\phi(a)$ to be this common label, assigning an arbitrary label in $[N]$ at each isolated vertex. Every selected edge has endpoints $\phi(f_{\text{L}}) = u_f$ and $\phi(f_{\text{R}}) = w_f$ and is present, so ϕ satisfies (2). Hence the source instance accepts.

The capacity $1 + e(P)N^2 \le n$ permits all the objects and the required padding. Only edge-check entries depend on the source input, each as $\overline{X}_{f,u,w}$; every other output bit is constant. Thus the construction is a depth-zero projection. Appending zero rows preserves every selected XOR, proving the result for all $m \ge m_{\oplus}(P, N)$. □

For even k, selecting k padding indices gives XOR zero in every row, so this construction fails whenever that many padding indices are available. The separate lift in Lemma D.1 gives the even-parameter lower bounds stated in Corollary D.2, with the fixed-parameter consequence in Equation (111).

3.3 k-SUM

We encode the objects in (7), together with padding to length n, as an indexed list of residues modulo 2^m. The target accepts exactly when the residues at k distinct indices sum to zero modulo 2^m (Definition 1.1); repeated residue values are allowed. We first represent each object by a signed integer vector whose zero-sum equations enforce the required choices. We then pack its coordinates into one residue, choosing a base large enough that a modular zero sum is equivalent to an exact vector zero sum. This construction works for either parity of k.

Role	coordinate	$\boldsymbol{\alpha^\star}$	c_1^{11}	c_1^{12}	$\boldsymbol{c_1^{21\star}}$	c_1^{22}	c_2^{11}	$\boldsymbol{c_2^{12\star}}$	c_2^{21}	c_2^{22}	δ_1	δ_2	δ_3
Selection	G_{f_1}	-1	1	1	1	1	0	0	0	0	0	0	0
Selection	G_{f_2}	-1	0	0	0	0	1	1	1	1	0	0	0
Agreement	C_{p_2,f_2}	0	-1	-2	-1	-2	1	1	2	2	0	0	0
Edges	Q	0	$\overline{X}^1_{11(1)}$	$\overline{X}^1_{12(1)}$	$\overline{X}^1_{21(0)}$	$\overline{X}^1_{22(1)}$	$\overline{X}^2_{11(1)}$	$\overline{X}^2_{12(0)}$	$\overline{X}^2_{21(0)}$	$\overline{X}^2_{22(1)}$	0	0	0
Padding	Q_{pad}	0	0	0	0	0	0	0	0	0	1	1	1

$$B = 16, \quad \Phi = G_{f_1} + BG_{f_2} + B^2 C_{p_2,f_2} + B^3 Q + B^4 Q_{\text{pad}}, \quad M = B^5 = 2^{20}.$$

standard residue	$\boldsymbol{\alpha^\star}$	c_1^{11}	c_1^{12}	$\boldsymbol{c_1^{21\star}}$	c_1^{22}	c_2^{11}	$\boldsymbol{c_2^{12\star}}$	c_2^{21}	c_2^{22}	δ_1	δ_2	δ_3
hexadecimal	FFFEF	00F01	00E01	FFF01	00E01	01110	00110	00210	01210	10000	10000	10000

Selected objects	exact signed sum y_S	sum modulo M (hexadecimal)
$\{\alpha, c_1^{21}, c_2^{12}\}$	$(0,0,0,0,0)$	00000
$\{\alpha, c_1^{21}, c_2^{21}\}$	$(0,0,1,0,0)$	00100
$\{\alpha, c_1^{11}, c_2^{12}\}$	$(0,0,0,1,0)$	01000
$\{\delta_1, \delta_2, \delta_3\}$	$(0,0,0,0,3)$	30000
$\{c_1^{11}, c_1^{12}, \delta_1\}$	$(2,0,-3,2,1)$	11D02

Figure 5: The five signed coordinates and twelve packed residues for the toy k-SUM instance, with $c_i^{uw} := c_{f_i,u,w}$ and padding indices $\delta_1, \delta_2, \delta_3$. Coordinate C_{p_2,f_2} subtracts the label proposed at p_2 by $f_1 = \rho(p_2)$ from the label proposed by f_2; coordinate Q counts absent selected edges. Gray entries are fixed nonzero constants; boxes display absence literals from Figure 2, with evaluated bits in parentheses. Stars mark the accepting selection. For this example, take base $B = 16$, let Φ be the displayed weighted sum of coordinates, and set modulus $M = B^5 = 2^{20}$. The lower tables give Φ modulo M in hexadecimal; for $m > 20$, multiply every residue by 2^{m-20}.

We use exact integer equations and positional packing, as in Karp's Exact Cover to Knapsack reduction [Kar72]. Represent the object at index j by a vector $\text{vec}(j) \in \mathbb{Z}^{R(P)}$, with $R(P)$ given by (30). We define its coordinates below and set all unspecified entries to zero. For a set $S \subseteq [n]$ of k

indices, write

$$y_S := \sum_{j \in S} \mathrm{vec}(j). \tag{28}$$

The upper matrix of Figure 5 shows these signed vectors as columns labeled by their objects; the lower tables show their packed residues and selected sums.

Exact selection and padding exclusion. For each $f \in E(P)$, introduce a selection coordinate G_f, giving $e(P)$ coordinates. It is -1 on α and 1 on each edge object in group f, so the intended choice of α and one object per group has sum $-1+1=0$ on every G_f. Add one padding coordinate Q_{pad}, equal to 1 on every padding object.

Let $\zeta \in \{0,1\}$ indicate whether S contains α, let c_f count its edge objects in group f, and let d count its padding objects. The sums on G_f and Q_{pad} are $c_f - \zeta$ and d, respectively. If $y_S = 0$, these give $c_f = \zeta$ and $d = 0$. Since $|S| = k$ and $e(P) = k-1$, we obtain $k = \zeta + \sum_f c_f + d = k\zeta$. Thus $\zeta = 1$ and $c_f = 1$ for every f: exactly the auxiliary object and one edge object per group are selected.

In Figure 5, the fourth selection has sum 3 on Q_{pad}, while the fifth has sum 2 on G_{f_1} and 1 on Q_{pad}; neither has vector sum zero.

Edge presence. A single coordinate Q checks all selected edges: set it to $\overline{X}_{f,u,w}$ on every edge object $c_{f,u,w}$. Once the selection equations force one object c_{f,u_f,w_f} per group, the sum on Q is $\sum_{f \in E(P)}(1 - X_{f,u_f,w_f})$. Every summand is nonnegative, so this sum is zero exactly when all represented edges are present. Unlike XOR, exact integer addition allows these checks to share one coordinate without cancellation.

Endpoint agreement. For each $(a,g) \in \mathcal{I}(P)$, introduce one comparison coordinate $C_{a,g}$, giving $|\mathcal{I}(P)|$ coordinates. Put $f = \rho(a)$ and assign $-\mathrm{lab}_a(f,u,w)$ on edge objects in group f and $+\mathrm{lab}_a(g,u,w)$ on edge objects in group g. Here we use the integer labels themselves. With one selected object per group, the sum on this coordinate is

$$\mathrm{lab}_a(g,u_g,w_g) - \mathrm{lab}_a(f,u_f,w_f), \tag{29}$$

which is zero exactly when the proposed labels agree. The reference comparisons therefore make all selected edges at a shared endpoint agree on its label.

For the running example, C_{p_2,f_2} is $-w$ on c_1^{uw} and $+u$ on c_2^{uw}, comparing the second index of the first edge with the first index of the second. In the accepting selection $\{\alpha, c_1^{21}, c_2^{12}\}$, this gives $0-1+1=0$. In the second selection, replacing c_2^{12} by c_2^{21} gives $0-1+2=1$, detecting the disagreement at p_2.

Number of coordinates. There are $e(P)$ selection coordinates, $|\mathcal{I}(P)|$ comparison coordinates, one edge-check coordinate and one padding coordinate. Thus

$$R(P) := e(P) + |\mathcal{I}(P)| + 2 \le 3e(P) + 2 = O(k), \tag{30}$$

by Equation (10).

Packing into a residue. Fix an order of the $R = R(P)$ coordinates. Every coordinate of a single vector has absolute value at most N, so every coordinate of y_S has absolute value at most kN. We give successive coordinates powers of a base $B > 2kN$. This separation will ensure that a nonzero vector sum cannot pack to zero modulo B^R. Set

$$q := \lceil \log(2kN + 1) \rceil, \qquad B := 2^q, \qquad m_\Sigma(P, N) := qR. \tag{31}$$

Then $B > 2kN$ and $2^{m_\Sigma(P,N)} = B^R$. For an integer vector $z = (z_0, \ldots, z_{R-1}) \in \mathbb{Z}^R$, define the integer packing and its standard residues by

$$\Phi(z) := \sum_{i=0}^{R-1} z_i B^i, \qquad \widehat{a}_j := \Phi(\mathrm{vec}(j)) \bmod B^R \in [0, B^R). \tag{32}$$

At the native width $m_\Sigma(P, N)$, the target list consists of the residues $\widehat{a}_j$. At any requested width $m \geq m_\Sigma(P, N)$, use

$$a_j := 2^{m - m_\Sigma(P,N)} \widehat{a}_j \in [0, 2^m). \tag{33}$$

Multiplication by this power of two appends $m - m_\Sigma(P, N)$ zero bits at the least significant end (the right) of each native encoding. The following lemma proves that packing and scaling preserve exactly the vector zero sums.

Lemma 3.1 (Packing preserves zero sums)**.** *For every set $S \subseteq [n]$ of k indices and every $m \geq m_\Sigma(P, N)$,*

$$\sum_{j \in S} a_j \equiv 0 \pmod{2^m} \iff y_S = 0. \tag{34}$$

Proof. We have $\|y_S\|_\infty \leq kN < B/2$. For any integer vector y in this range, suppose $B^R \mid \Phi(y)$. Reducing modulo B gives $B \mid y_0$, and $|y_0| < B/2$ forces $y_0 = 0$. Dividing by B leaves $B^{R-1} \mid \sum_{i=1}^{R-1} y_i B^{i-1}$, so the same argument forces $y_1 = 0$. Repeating gives $y = 0$. The converse is immediate, so

$$\Phi(y) \equiv 0 \pmod{B^R} \iff y = 0 \qquad (\|y\|_\infty \leq kN). \tag{35}$$

Linearity gives $\sum_{j \in S} \widehat{a}_j \equiv \Phi(y_S) \pmod{B^R}$, proving the claim at the native width. For any integer Z, cancelling the common power of two gives

$$2^m \mid 2^{m - m_\Sigma(P,N)} Z \iff 2^{m_\Sigma(P,N)} \mid Z. \tag{36}$$

Applying this with $Z = \sum_{j \in S} \widehat{a}_j$ proves the claim for every larger width as well. □

Proof of Theorem 2.3(iii). Let $\phi : V(P) \to [N]$ be an accepting source map as in Definition 2.1. Select the indices of α and $c_{f,\phi(f_\mathrm{L}),\phi(f_\mathrm{R})}$ for every $f \in E(P)$. These k distinct indices contain no padding and have $y_S = 0$ by the selection equations, edge presence and endpoint agreement. Hence Lemma 3.1 makes the constructed SUM instance accept.

Conversely, let S be a set of k indices whose residues sum to zero modulo 2^m. By Lemma 3.1, $y_S = 0$. The exact selection equations force α and one edge object c_{f,u_f,w_f} per group, with no padding. Coordinate Q forces $X_{f,u_f,w_f} = 1$ for every f, and the comparison coordinates force endpoint agreement. Define $\phi(a)$ to be the common label at each non-isolated vertex, assigning arbitrary labels in $[N]$ to isolated vertices. Then $\phi(f_\mathrm{L}) = u_f$ and $\phi(f_\mathrm{R}) = w_f$ for every f, so edge presence gives (2), and the source accepts.

The capacity $1 + e(P)N^2 \leq n$ permits the auxiliary object, all edge objects and the required padding. The auxiliary and padding encodings are constant. Each edge object's vector depends

only on its own source bit $X_{f,u,w}$ through coordinate Q. Its packed and scaled residue therefore takes one of two fixed m-bit values, according as $X_{f,u,w} = 0$ or 1. Each output bit is consequently a constant, $X_{f,u,w}$ or $\overline{X}_{f,u,w}$, so the construction is a depth-zero projection, including any carries in the packing. Finally, Equations (30) and (31) give $m_\Sigma(P, N) = O(k \log(2kN))$. □

Larger budgets. Every larger budget is allowed: append zero coordinates for $D \geq D_0(P, N, n)$ in OV, append zero rows for $m \geq m_\oplus(P, N)$ in XOR, and use Equations (33) and (36) for $m \geq m_\Sigma(P, N)$ in SUM.

4 Unconditional Lower Bounds

We now combine the projections of Theorem 2.3 with four source lower bounds. For fixed k, the LRR theorem gives an exponent linear in k. For growing k, Beame's clique bound applies at every fixed depth; direct counting at depth two and a minterm argument for Σ_3 give exponents linear in k within their stated polynomial ranges. Each application checks that the hosts in Lemma 2.4 meet the source theorem's requirements.

4.1 Fixed parameters

For fixed k, the expander pattern is fixed as well, so the LRR theorem applies once the host is sufficiently large. Its exponent grows linearly with k along our pattern family.

4.1.1 Fixed-pattern hardness

Let $\kappa(P)$ denote the colored parameter $\kappa_{\mathrm{col}}(P)$ of [LRR17, Def. 2.12(iii)], following the notation of [Ros21, Footnote 1]. This parameter determines the exponent in the following lower bound.

Theorem 4.1 (Li–Razborov–Rossman fixed-pattern lower bound [LRR17])**.** *For every fixed simple loopless pattern P with at least one edge and no isolated vertices, every fixed depth d, and every constant $\eta > 0$, there is a threshold $N_0(P, d, \eta)$ such that*

$$\mathrm{size}_d(P\text{-SUB}_N) + 2e(P)N^2 + 2 \geq N^{\kappa(P)-\eta} \tag{37}$$

for every $N \geq N_0(P, d, \eta)$. If $\kappa(P) - \eta > 2$, then the threshold can be chosen so that the additive term accounting for input gates is absorbed and $\mathrm{size}_d(P\text{-SUB}_N) \geq N^{\kappa(P)-\eta}$.

Appendix A.2 derives this statement in our gate-count convention and explains how a small loss in the exponent absorbs the additive term. The theorem applies separately to each fixed pattern and does not control the required host size as P grows.

For the common family of Lemma 2.2, the construction in Appendix B.1 gives

$$\kappa(P_k^{\mathrm{exp}}) \geq c_\kappa(k-1), \qquad c_\kappa := 1/128. \tag{38}$$

The threshold $k_0 \geq 600$ fixed in Lemma 2.2 ensures $c_\kappa(k_0 - 1) > 4$, leaving room to absorb the input-gate term.

Corollary 4.2 (Fixed-k source bound with a universal exponent constant)**.** *Fix any constant γ with $0 < \gamma < c_\kappa$. For every fixed depth d and every fixed integer $k \geq k_0$, there is a threshold $N_0(\gamma, d, k)$ such that*

$$\mathrm{size}_d(P_k^{\mathrm{exp}}\text{-SUB}_N) \geq N^{\gamma(k-1)} \tag{39}$$

for every $N \geq N_0(\gamma, d, k)$.

Proof. For fixed d and $k \geq k_0$, choose the exponent loss $\eta := (c_\kappa - \gamma)(k-1)/2 > 0$. By Lemma 2.2, the fixed pattern $P_k^{\exp}$ is simple and loopless, has $k-1 \geq 1$ edges, and has no isolated vertices, so Theorem 4.1 applies. Moreover, Equation (38) gives

$$\kappa(P_k^{\exp}) - \eta \geq \frac{c_\kappa + \gamma}{2}(k-1) \geq \tfrac{1}{2} c_\kappa (k_0 - 1) > 2.$$

Since $(c_\kappa + \gamma)/2 > \gamma$, the final claim of Theorem 4.1 yields Equation (39) for all sufficiently large N. □

4.1.2 Transfer to the three targets

Applying Lemma 2.4 to Corollary 4.2 gives base n for OV and n/k for XOR and SUM. The same constant in the target exponent works at every fixed depth and for every sufficiently large fixed k; only the minimum required n may depend on (d, k).

Theorem 4.3 (Unconditional lower bounds for fixed k)**.** *Let $k_0 \in \mathbb{N}$ be the universal threshold of Lemma 2.2, and let $C_{\mathrm{OV}}, C_\oplus, C_\Sigma$ be the universal construction constants fixed in Theorem 2.3 and Lemma 2.4. There is a universal constant $\beta > 0$, independent of k, d, and n, such that for every fixed depth d and every fixed integer $k \geq k_0$ there is a threshold $n_0(d,k)$ for which the following hold for every $n \geq n_0(d,k)$:*

$$\mathrm{size}_d(k\text{-}\mathsf{OV}_{n,D,k}) \geq n^{\beta(k-1)} \qquad \text{for every } D \geq C_{\mathrm{OV}} k \log n, \tag{40}$$

$$\mathrm{size}_d(k\text{-}\mathsf{XOR}_{n,m,k}) \geq (n/k)^{\beta(k-1)} \qquad \text{for odd } k \text{ and every } m \geq C_\oplus k \log(en/k), \tag{41}$$

$$\mathrm{size}_d(k\text{-}\mathsf{SUM}_{n,m,k}) \geq (n/k)^{\beta(k-1)} \qquad \text{for every } m \geq C_\Sigma k \log(en/k). \tag{42}$$

Proof. Choose universal constants

$$0 < 2\beta < \gamma < c_\kappa. \tag{43}$$

For each fixed (d, k), Corollary 4.2 gives

$$\mathrm{size}_d(P_k^{\exp}\text{-SUB}_N) \geq N^{\gamma(k-1)} \qquad (N \geq N_0(\gamma, d, k)). \tag{44}$$

Choose $n_0(d,k)$ so that $n \geq \max\{16k, k^3\}$, both hosts in Equation (11) are at least $N_0(\gamma, d, k)$, and $(n/k)^{\gamma/2-\beta} \geq 2^\gamma$. All conditions hold eventually for fixed (d, k). Apply Lemma 2.4 with source exponent $a = \gamma(k-1)$ and target rate $c = \beta$, using the sharper SUM budget (18). The condition (20) holds for both bases and gives the three bounds, with oddness required only for k-XOR. □

For certain even witness sizes, the lift in Lemma D.1 transfers the odd-k XOR bound (41). Fixed tag rows force target witnesses to select whole groups of copies, encoding odd-size source witnesses. The resulting exponent is controlled by the odd source witness size; Corollary D.2(i) gives the parameter conditions and quantitative bound.

Theorems C.1 and C.2 extend the list bounds to general row and bit-width budgets. At principal budgets, Table 4 compares the lower bounds with elementary Σ_3 upper bounds, including the SUM block-carry construction; the upper-bound proofs are in Appendix E.

4.2 A clique bound at every fixed depth[1]

For growing k, we need a source theorem that explicitly relates the host size to the pattern size. Beame's small-clique bound provides this relation. A clique with at most $k-1$ edges has only $O(\sqrt{k})$

[1]The use of Beame's small-clique lower bound [Bea90] as the source statement for this subsection was suggested by the AI language model GPT-5.5; the authors independently verified its applicability and developed the resulting gate-count consequence (Lemma 4.4) and Theorem 4.5.

vertices, limiting the dependence on k in Beame's exponent to order $\sqrt{k}$. Its additional requirement of at most $\log N$ clique vertices limits the exponent further once k reaches order $\log^2 n$.

Lemma 4.4 (Gate-count consequence of Beame's small-clique lower bound)**.** *For every fixed depth d, there are a constant $a_d > 0$ and integers $r_d, N_d \geq 1$ such that, whenever $N \geq N_d$ and $r_d \leq r \leq \log N$,*

$$\mathrm{size}_d(K_r\text{-CLIQUE}_N) \geq N^{a_d r}. \tag{45}$$

Here K_r-CLIQUE$_N$ is the function on $\binom{N}{2}$ edge variables that accepts exactly when the graph contains an r-vertex clique (Definition B.3). The upper limit $r \leq \log N$ is independent of depth; only a_d, r_d, N_d may depend on d. The proof permits $a_d := 1/(2000(d+2)^3)$ and $r_d \geq 3000(d+2)^3$, after enlarging r_d to absorb Beame's absolute threshold for r.

Appendix A.3 derives this circuit bound from Beame's PRIORITY CRCW PRAM memory-cell bound, including the simulation and the constants needed at small depths.

The explicit range $r \leq \log N$ is essential here. Rossman's stronger exponent is proved for every fixed number r of clique vertices; a growing-r extension is presented only as a possibility [Ros08, Thm. 1.2; Sec. 6, final paragraph before "Open Questions"]. That theorem therefore does not control the required host size for the choice $r = r(n)$ used below.

For $r \geq 2$ and $t := k - 1 - \binom{r}{2} \geq 0$, use the padded clique $P^{\mathrm{clq}}_{k,r} := K_r \dot\cup J_t$ of Definition B.3, where J_t is a matching of t edges. Lemma B.4 projects ordinary clique to K_r-SUB$_N$, then to the padded pattern by fixing all matching-edge variables to 1. Thus Beame's bound applies to a source with exactly the $k-1$ edges required by Theorem 2.3.

Theorem 4.5 (Growing k at every fixed depth)**.** *Fix a constant $\varepsilon > 0$. For every fixed depth d, the following hold along every sequence $k = k(n)$ satisfying $k \to \infty$ and $k \leq n^{1-\varepsilon}$, for all sufficiently large n. Define*

$$\rho_{\mathrm{OV}} := \min\{\sqrt{k-1}, \log n\}, \qquad \rho := \min\{\sqrt{k-1}, \log(n/k)\}. \tag{46}$$

Then

$$\mathrm{size}_d(k\text{-}\mathsf{OV}_{n,D,k}) \geq \exp\bigl(\Omega_d(\rho_{\mathrm{OV}} \log n)\bigr) \qquad \text{for every } D \geq C_{\mathrm{OV}} k \log n, \tag{47}$$

$$\mathrm{size}_d(k\text{-}\mathsf{XOR}_{n,m,k}) \geq \exp\bigl(\Omega_d(\rho \log(n/k))\bigr) \qquad \text{for odd } k \text{ and every } m \geq C_{\oplus} k \log(en/k), \tag{48}$$

$$\mathrm{size}_d(k\text{-}\mathsf{SUM}_{n,m,k}) \geq \exp\bigl(\Omega_d(\rho \log(n/k))\bigr) \qquad \text{for every } m \geq C_{\Sigma} k \log n. \tag{49}$$

The implied constants in the three lower bounds depend only on d.

Proof. Fix d and the constants a_d, r_d, N_d from Lemma 4.4. Let Q be one of k-$\mathsf{OV}_{n,D,k}$, k-$\mathsf{XOR}_{n,m,k}$, or k-$\mathsf{SUM}_{n,m,k}$. Use N_Q, B_Q from Lemma 2.4 and let ρ_Q be ρ_{OV} for OV and ρ otherwise. Since $k \to \infty$ and $k \leq n^{1-\varepsilon}$, eventually $n \geq 16k$, $N_Q \geq N_d$ and $\rho_Q \to \infty$. Choose

$$r := \lfloor \rho_Q/4 \rfloor . \tag{50}$$

For sufficiently large n, $r \geq r_d$ and $r \geq \rho_Q/8$. The clique fits the edge allowance, since $r \leq \sqrt{k-1}$ gives $\binom{r}{2} \leq r^2/2 \leq k-1$. Beame's upper cap holds as well: $\rho_Q \leq \log B_Q$ gives $r \leq \frac{1}{4}\log B_Q \leq \log N_Q$ by Lemma 2.4. Thus the padded pattern of Definition B.3 has exactly $k-1$ edges and satisfies

$$\mathrm{size}_d(P^{\mathrm{clq}}_{k,r}\text{-SUB}_{N_Q}) \geq \mathrm{size}_d(K_r\text{-CLIQUE}_{N_Q}) \geq N_Q^{a_d r}, \tag{51}$$

by Lemmas B.4 and 4.4. Applying Equation (19) gives $\log \mathrm{size}_d(Q) \geq (a_d/32)\rho_Q \log B_Q$, with constants depending only on d.

For the budgets, use the general SUM branch of Lemma 2.4: the full range $k \leq n^{1-\varepsilon}$ need not satisfy $n \geq k^3$. The OV and odd-XOR budgets are unchanged. The clique's growing degree costs nothing extra, since Equation (10) bounds the number of endpoint comparisons by $2e(P)$, independently of $\Delta(P)$. □

When increasing k no longer improves the bound. At the budgets of Theorem 4.5, the inequalities $\log(n/k) \geq \varepsilon \log n$ and $\sqrt{k-1} \geq \sqrt{k}/2$ give, for all three targets with odd k for k-XOR,

$$\log \mathrm{size}_d \geq \Omega_{d,\varepsilon}\big(\min\{\sqrt{k}, \log n\} \cdot \log n\big). \tag{52}$$

If $k \to \infty$ and $k = n^{o(1)}$, then $\log(n/k) = (1 - o(1)) \log n$, so for all three targets, with odd k for k-XOR,

$$\mathrm{size}_d \geq n^{\Omega_d(\min\{\sqrt{k}, \log n\})}. \tag{53}$$

In the subpolynomial range of Equation (53), once $k \geq \log^2 n$, the logarithmic limit on clique vertices fixes the bound at $\exp(\Omega_d(\log^2 n))$.

4.3 Depth two: counting terms and clauses

At depth two, we can count terms and clauses directly. Every colored copy requires a distinct DNF term. For CNFs, we count minimal sets of host edges meeting every colored copy: each such set requires a distinct clause. We use $P_k^{\exp}$ from Lemma 2.2 because it has no isolated vertices, has $\Theta(k)$ vertices, and contains a vertex of degree at least two; the proof does not use expansion or κ.

For a map $\phi : V(P) \to [N]$, let

$$E_\phi := \{X_{f,\phi(f_{\mathrm{L}}),\phi(f_{\mathrm{R}})} : f = (f_{\mathrm{L}}, f_{\mathrm{R}}) \in E(P)\}. \tag{54}$$

The assignment with positive support exactly E_ϕ is a minimal positive input of P-SUB$_N$, and

$$P\text{-SUB}_N = \bigvee_{\phi : V(P) \to [N]} \bigwedge_{x \in E_\phi} x. \tag{55}$$

Theorem 4.6 (Depth two, both orientations)**.** *There are universal constants $k_D \geq k_0$, $\gamma_D > 0$, and $\xi_D > 0$ such that the following hold.*

(i) For every $k \geq k_D$ and every $N \geq k^2$,

$$\mathrm{size}_2(P_k^{\exp}\text{-SUB}_N) \geq N^{\gamma_D k}. \tag{56}$$

(ii) For every $k \geq k_D$, the following target bounds hold whenever the indicated range condition is satisfied:

(a) if $k \leq \xi_D n^{1/4}$, then

$$\mathrm{size}_2(k\text{-OV}_{n,D,k}) \geq n^{\Omega(k)} \tag{57}$$

for every $D \geq C_{\mathrm{OV}} k \log n$;

(b) if $k \leq \xi_D n^{1/5}$, then

$$\mathrm{size}_2(k\text{-XOR}_{n,m,k}) \geq (n/k)^{\Omega(k)} \quad \text{for odd } k \text{ and every } m \geq C_{\oplus} k \log(en/k), \tag{58}$$

$$\mathrm{size}_2(k\text{-SUM}_{n,m,k}) \geq (n/k)^{\Omega(k)} \quad \text{for every } m \geq C_{\Sigma} k \log(en/k). \tag{59}$$

All implied constants are universal.

Proof. Fix $P = P_k^{\exp}$. We first lower-bound both possible top-gate orientations of a depth-two source circuit.

DNF orientation. Let $\mathcal{D}$ be a DNF, with arbitrary input negations, computing P-SUB$_N$. We show that each term can accept at most one minimal positive input. For every ϕ, the minimal positive input E_ϕ satisfies some term T_ϕ of $\mathcal{D}$. Its positive literals all lie in E_ϕ, and its negative literals involve only variables outside E_ϕ. Setting exactly those positive variables to 1 and all others to 0 still satisfies the term. If a variable $x \in E_\phi$ were omitted, this assignment would have no positive variable in the pattern-edge group containing x, so it could not contain a colored copy. Correctness therefore forces the positive part of T_ϕ to be exactly E_ϕ.

If the same term were satisfied by another minimal positive input $E_{\phi'}$, the same argument would force its positive part to equal $E_{\phi'}$ as well. Hence $E_{\phi'} = E_\phi$, so each term accepts at most one minimal positive input. Distinct maps give distinct colored edge sets: if $\phi' \neq \phi$, choose a vertex a on which they differ; because P has no isolated vertices, an edge incident to a contributes different input coordinates to E_ϕ and $E_{\phi'}$. Since every E_ϕ must satisfy at least one term, these inputs require distinct terms, giving at least

$$N^{v(P)} = N^{\Omega(k)} \tag{60}$$

terms, where the maximum-degree bound gives $v(P) \geq 2e(P)/4 = (k-1)/2$.

CNF orientation. By Equation (55), the source function is monotone. Let $\mathcal{C}$ be a CNF computing it, and discard tautological clauses, which contain both a variable and its negation and are therefore always true. For each remaining clause C, let P_C and N_C be the sets of variables occurring positively and negatively, respectively; these sets are disjoint. Every 1-input sets some variable of P_C to 1: otherwise, setting all variables of N_C to 1 would preserve acceptance by monotonicity while falsifying C. Every 0-input still falsifies some positive part, because a clause falsified on that input has all its positive literals equal to 0. Thus deleting all negative literals gives an equivalent monotone CNF without increasing its clause count.

In any monotone CNF computing the function, every clause is a *transversal*: a set meeting every E_ϕ in (54). A clause disjoint from some E_ϕ would be falsified by the 1-input with positive support E_ϕ. Moreover, every minimal transversal must occur as a clause. To see this, let W be minimal and set the variables in W to 0 and all others to 1. Since W meets every colored copy, this is a 0-input, so some clause C is falsified and has only positive literals, hence satisfies $C \subseteq W$. The clause C is itself a transversal, and minimality of W forces $C = W$.

To construct many minimal transversals, think of deleting edges from the complete colored host. A transversal is a set of deletions that destroys every colored copy, and it is minimal if restoring any one deleted edge creates a copy. Each such set forces a clause saying that at least one of those edges must be present.

Fix a pattern vertex a of degree $\delta \geq 2$. Every copy assigns some label i to a and uses an edge in every incident pattern-edge group. To rule out copies using label i, choose one incident group and delete all its host edges with label i at a. We choose which incident group to delete from independently for each of the N labels, giving δ^N ways to specify the deletions.

Let $E_P(a)$ be the incident pattern edges and let $h : [N] \to E_P(a)$ specify the edge group chosen for each label. In the notation of (8), W_h is the set of variables representing the deleted host edges:

$$W_h := \{X_{f,u,w} : \text{there is } i \in [N] \text{ with } f = h(i) \text{ and } \mathrm{lab}_a(f,u,w) = i\}. \tag{61}$$

Setting the variables in W_h to 0 and all other host-edge variables to 1 gives a rejecting input: a copy with $\phi(a) = i$ would need an edge in group $h(i)$ with label i at a, and every such edge has been deleted. Thus W_h is a transversal.

Now restore any deleted edge $x \in W_h$, say in group $h(i) = \{a, b\}$. Assign a the label i, assign b the other endpoint label of x, and choose all remaining labels arbitrarily. Every edge of this copy

is present: x was restored, its other edges incident to a lie in groups other than $h(i)$, and edges not incident to a were never deleted. Hence restoring any deleted edge creates a copy, so W_h is minimal.

Different functions h give different deletion sets: if $h(i) \neq h'(i)$, the host edges in group $h(i)$ with label i at a are deleted for W_h and remain present for $W_{h'}$. Hence the δ^N assignments of incident groups to labels give distinct minimal transversals, so every CNF has at least

$$\delta^N \geq 2^N \tag{62}$$

clauses. Since $N \geq k^2 \geq 16$, we have $\log N \leq \sqrt{N} \leq N/k$, so $2^N \geq N^k$.

To turn these counts into gate counts, first propagate constants. Every gate feeding the output of a depth-at-most-two circuit then has only literal inputs. For an OR output, bypass any OR gate feeding it by wiring that gate's literal inputs directly to the output; ORing these inputs gives the same result. What remains represents a DNF whose terms come from AND gates or from literals wired directly to the output. For an AND output, bypassing AND gates feeding it similarly gives a circuit representing a CNF, with OR gates supplying clauses and direct literals supplying one-literal clauses. Here DNF and CNF refer to Boolean expressions that allow terms or clauses consisting of a single literal; for example, $(x \wedge y) \vee z$ is a DNF with term z. The direct wires represent these terms or clauses without inserting or counting any additional gates.

Every term counted in the DNF argument contains $k - 1 \geq 2$ positive literals, so it comes from an AND gate rather than a direct literal input. Similarly, each required clause W_h contains $N^2 \geq 2$ variables: each of the N labels at a contributes N variables, one for each label at the opposite endpoint. Thus it comes from an OR gate. Each such gate has a fixed set of literal inputs and therefore supplies only one term or clause; deleting negative literals only shortens an existing clause. Hence distinct counted terms or clauses require distinct non-input gates. Choosing $0 < \gamma_D \leq 1/3$ makes both gate counts at least $N^{\gamma_D k}$, proving Equation (56).

Transfer to the targets. Choose $\xi_D \leq 1/2$. The OV range $k \leq \xi_D n^{1/4}$ and list range $k \leq \xi_D n^{1/5}$ ensure $n \geq 16k$, as required by Lemma 2.4. The list range also ensures $n \geq k^3$, which permits the sharper SUM budget in Equation (18). The lemma gives $N_{\mathrm{OV}} \geq \sqrt{n}/2 \geq k^2$ in the OV range and $N_{\oplus,\Sigma} \geq \sqrt{n/k}/2 \geq k^2$ in the list range. Thus both hosts satisfy the source threshold $N \geq k^2$ in Equation (56). Apply Equation (19) with $a = \gamma_D k$ to obtain exponents at least $\gamma_D k/4$ in the respective bases. The projections preserve depth and both orientations, with oddness needed only for k-XOR. □

Subpolynomial parameters. If $k \to \infty$ and $k = n^{o(1)}$, both explicit ranges of Theorem 4.6 eventually hold and $n/k = n^{1-o(1)}$. At the theorem's budgets, all three target bounds therefore become

$$\mathrm{size}_2 \geq n^{\Omega(k)} = \exp\bigl(\Omega(k \log n)\bigr), \tag{63}$$

with odd k for k-XOR. At $k = \Theta(\log^2 n)$ this is $\exp(\Omega(\log^3 n))$, stronger than the clique bound (53).

Appendix D gives a sharper, tight depth-two bound for odd XOR through a PARITY projection, together with even-parameter consequences.

4.4 Depth three: minterms and a top disjunction[2]

At depth three with a top disjunction, we strengthen the growing-k bounds of Theorem 4.5 to exponents linear in k for all three targets in the polynomial range stated below. The projections and host substitutions remain those of Theorem 2.3 and Lemma 2.4; the new ingredient is a probability bound for one CNF. We state that bound and its circuit consequence here, with the complete proofs in Appendix F.

A Σ_3 circuit is a disjunction of CNFs. On an accepting input, one middle CNF must accept; on a rejecting input, every middle CNF must reject. We exploit this by adding a random colored copy to a sparse background, with parameters that, with high probability, make the planted copy unique and keep its edges out of the background. On this event, the full planting accepts, while deleting any planted edge makes the source reject. Any middle CNF accepting the full planting must reject every proper deletion as well. The key bound shows that a fixed CNF is unlikely to have this property, so a correct circuit needs many middle CNFs.

For the probabilistic statement, fix a simple loopless pattern F with endpoint orders as in Definition 2.1, a host size $N \geq 1$, and a density $p \in (0, 1]$. An input is represented by its set of present host-edge variables; variables outside the set have value 0. Sample a background G by including each variable independently with probability p. Independently of G, choose the labels $\phi^*(v) \in [N]$ independently and uniformly for all pattern vertices, obtaining the planted copy

$$F^* := \{X_{f,\phi^*(f_{\mathrm{L}}),\phi^*(f_{\mathrm{R}})} : f = (f_{\mathrm{L}}, f_{\mathrm{R}}) \in E(F)\}. \tag{64}$$

For a fixed CNF T, let $\mathcal{M}(T)$ be the *full minterm event*: $T(G \cup F^*) = 1$ while $T(G \cup R) = 0$ for every proper subset $R \subsetneq F^*$. Thus every planted edge is needed for T to accept over the same fixed background.

Theorem 4.7 (Single-CNF minterm bound)**.** *Let $I \subseteq V(F)$ be a nonempty independent set of L positive-degree vertices. With G and F^* sampled as above, there is an absolute constant $C_0 > 0$ such that every fixed CNF T with $s := s(T) \geq 1$ clauses satisfies*

$$\Pr[\mathcal{M}(T)] \leq \left(\frac{C_0\big(L + \log(\mathrm{e}s)\big)}{pN}\right)^L. \tag{65}$$

The probability is over G and the planted labels. Clause widths and literal signs are unrestricted, and T need not compute the source function. If $s = 0$, then $T \equiv 1$ and the event is empty.

The Lth power comes from revealing the labels at the L vertices of I, which are independent and uniform. Once the outside labels are fixed, each such label determines all planted edges incident to its vertex, since all its neighbors lie outside I. A false clause can become true when one of those edges sets a positive variable to 1. If many candidate labels can do this, the random background is likely to satisfy the clause already; if only a few can, a uniform planted label is unlikely to select one.

A union bound over the s clauses and the 2^L possible sets of revealed vertices puts the cutoff between many and few at order $(L+\log(\mathrm{e}s))/p$, accounting for the numerator in (65). Appendix F.1 formalizes the argument by revealing the labels in I one at a time, each together with the planted edges at its vertex. No restriction on clause width is used; Section 1.3 compares this with the width-dependent bound of Choudhury and Sreenivasaiah [CS26].

[2]The AI language model GPT-5.6 Sol suggested combining an adaptive process for revealing planted labels with a planted-copy minterm formulation. The authors independently verified the suggestion and developed the argument for the padded expander family; the AI disclosure records the antecedents and assistance in detail.

Applying the bound to the bipartite 4-regular core $F_k^{\exp}$ of Lemma 2.2 gives the source lower bound below. Padding restores the common $(k-1)$-edge pattern, and Lemma 2.4 transfers this lower bound to the three targets.

Theorem 4.8 (Depth three, top disjunction, exponent linear in k)**.** *For every constant γ with $0 < \gamma < 1/8$, there are $k_E(\gamma) \geq k_0$ and $C_E(\gamma) > 0$ such that*

$$\mathrm{size}_{\Sigma_3}(P_k^{\exp}\text{-SUB}_N) \geq \mathrm{size}_{\Sigma_3}(F_k^{\exp}\text{-SUB}_N) > N^{\gamma k} \tag{66}$$

for every $k \geq k_E(\gamma)$ and every $N \geq k^{C_E(\gamma)}$. For every constant c with $0 < c < \gamma/2$, there are $k_1(\gamma, c) \in \mathbb{N}$ and $C'(\gamma, c) > 1$ such that, for every $k \geq k_1(\gamma, c)$ and every $n \geq k^{C'(\gamma,c)}$,

$$\mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{OV}_{n,D,k}) \geq n^{ck} \qquad \text{for every } D \geq C_{\mathrm{OV}} k \log n, \tag{67}$$

$$\mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{XOR}_{n,m,k}) \geq (n/k)^{ck} \qquad \text{for odd } k \text{ and every } m \geq C_{\oplus} k \log(en/k), \tag{68}$$

$$\mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{SUM}_{n,m,k}) \geq (n/k)^{ck} \qquad \text{for every } m \geq C_{\Sigma} k \log(en/k). \tag{69}$$

Every fixed $0 < c < 1/16$ is therefore attainable.

Proof overview. Take I to be one bipartition side of the core; it has $L = k/4 + O(1)$ vertices. For $p = N^{-\theta}$ with θ just above $1/2$ and N a sufficiently large power of k, the planted copy is unique and none of its edges lies in the background with probability $1 - o(1)$ (Appendix F.2). The threshold $1/2$ comes from 4-regularity: changing the labels of j core vertices requires at least $2j$ background edges. With the planted labels fixed, each fixed set of j changed vertices therefore contributes at most $N^j p^{2j} = N^{-(2\theta-1)j}$ to the expected number of alternative copies. The polynomial host threshold also controls the number of possible changed sets.

On this event, correctness forces an accepting middle CNF to satisfy $\mathcal{M}(T)$. The main factor in (65) is $N^{-(1-\theta)L}$; for the clause counts arising from a circuit of size at most $N^{\gamma k}$, the polynomial host absorbs the numerator with an arbitrarily small loss in the exponent. Since $(1-\theta)L/k$ can approach $1/8$, a union bound rules out source circuits of size at most $N^{\gamma k}$ for every fixed $\gamma < 1/8$.

The resulting source lower bound is exactly what the common transfer lemma needs. Its square-root hosts convert the source coefficient γ to any target coefficient $c < \gamma/2$, giving every $c < 1/16$. Thus the additional work is confined to the source bound; the same projections give OV, odd-k XOR, and SUM at the stated budgets. Appendix F.3 supplies the parameter choices, padding argument, and threshold checks.

Corollary D.2(ii) combines Theorem 4.8 with the even-K lift; because the lift is a projection between k-XOR instances, it preserves the circuit orientation. In the subpolynomial regime $k \to \infty$ with $k = n^{o(1)}$, the polynomial relation in Theorem 4.8 eventually holds and $n/k = n^{1-o(1)}$, so the three bounds have the common form $n^{\Omega(k)} = \exp(\Omega(k \log n))$, with odd k in the direct k-XOR statement.

In the polynomial-host range of Theorem 4.8, take dimension $D = \Theta(k \log n)$ for OV and width $m = \Theta(k \log(en/k))$ for XOR and SUM, with constant factors large enough to satisfy its budget thresholds. Then Proposition E.1 gives OV and XOR circuits of sizes $n^{k+o(k)}$ and $(n/k)^{k+o(k)}$, while Proposition E.3 gives a SUM circuit of size $(n/k)^{O(k)}$, with odd k in the direct k-XOR comparison. These Σ_3 upper bounds match the theorem's lower bounds up to constant factors in the exponent.

The proof uses the top disjunction to select one accepting middle CNF, so Theorem 4.8 bounds only size_{Σ_3}. Since $\mathrm{size}_3 = \min\{\mathrm{size}_{\Sigma_3}, \mathrm{size}_{\Pi_3}\}$ (Equation (5)), unrestricted depth three also requires an analogous bound for the Π_3 orientation, which remains open (Section 5.2).

Comparison with PARITY. Appendix D.3 compares a supplementary PARITY projection for odd XOR with the subgraph bounds at the common row budget $m = \Theta(k \log(n/k))$, with a sufficiently large constant factor. At $k = \Theta(\log^2 n)$, PARITY gives the larger unconditional bound at depth two and for Π_3, while the subgraph bounds are larger for Σ_3 and fixed $d \geq 4$ (Table 3). The comparison at other growing witness sizes depends on both k and d.

5 Conditional Bounds and Open Problems

For the remaining depths and orientations, a stronger source theorem would give growing-k bounds with exponent linear in k through the same projections. We state this hypothesis, compare the targets, and identify the open questions.

5.1 A conditional consequence

The fixed-pattern theorem (Theorem 4.1) allows the minimum host size $N_0(P, d, \eta)$ to depend arbitrarily on P. For the padded expander family of Lemma 2.2, we ask for an exponent linear in k once the host size is polynomial in k, with constants depending only on the fixed depth.

Conjecture 5.1 (Pattern-Uniform LRR, depth-by-depth form)**.** *For the padded expander family $(P_k^{\mathrm{exp}})_{k \geq k_0}$, for every fixed depth d there are constants $\gamma_d > 0$ and $c'_d > 0$ such that*

$$\mathrm{size}_d(P_k^{\mathrm{exp}}\text{-SUB}_N) \geq N^{\gamma_d k} \tag{70}$$

for every $k \geq k_0$ and every $N \geq k^{c'_d}$.

Theorem 4.6(i) proves the depth-two case with $\gamma_2 = \gamma_D$ and $c'_2 = 2$, after replacing k_0 by k_D. For each fixed $0 < \gamma < 1/8$, Theorem 4.8 proves the Σ_3 version with threshold exponent $C_E(\gamma)$, after replacing k_0 by $k_E(\gamma)$. The Π_3 case, and hence the full $d = 3$ conjecture, remains open, as do all fixed $d \geq 4$ cases.

The switching estimate in [LRR17] does not directly give this polynomial threshold. With the restriction parameter $\delta > 0$ in Equation (80) fixed independently of k, this estimate certifies a source rate $\gamma > 0$ only when $\log N = \Omega_{d,\delta,\gamma}(k)$ (Lemma A.2). The scope discussion limits this observation to that estimate. The published range $d = o(\log N / \log\log N)$ for each fixed pattern [LRR17, p. 938] concerns growing depth and does not give the required uniformity over patterns.

A stronger rate. Requiring $\gamma_d \geq \gamma_0 > 0$ for one universal constant γ_0 would make the constants hidden by Ω_d below independent of depth, as in Theorem 4.3. The stated depth-by-depth form already suffices for the following consequence.

Corollary 5.2 (Conditional lower bounds at every fixed depth)**.** *Assume Conjecture 5.1. For every fixed depth d there are integers $k_F(d) \geq k_0$ and constants $C_F(d) > 1$ such that, for every $k \geq k_F(d)$ and every $n \geq k^{C_F(d)}$,*

$$\mathrm{size}_d(k\text{-}\mathsf{OV}_{n,D,k}) \geq \exp\bigl(\Omega_d(k \log n)\bigr) \qquad \text{for every } D \geq C_{\mathrm{OV}} k \log n, \tag{71}$$

$$\mathrm{size}_d(k\text{-}\mathsf{XOR}_{n,m,k}) \geq \exp\bigl(\Omega_d(k \log n)\bigr) \qquad \text{for odd } k \text{ and every } m \geq C_{\oplus} k \log(en/k), \tag{72}$$

$$\mathrm{size}_d(k\text{-}\mathsf{SUM}_{n,m,k}) \geq \exp\bigl(\Omega_d(k \log n)\bigr) \qquad \text{for every } m \geq C_{\Sigma} k \log(en/k). \tag{73}$$

Consequently, the conclusion applies along every sequence $k = k(n) \to \infty$ with $k = n^{o(1)}$. No new target reduction is used: the projections are exactly those of Theorem 2.3.

Proof. Fix d and the conjectural constants, and choose

$$C_F(d) \geq \max\{2c'_d + 4, 4\}. \tag{74}$$

Take $k_F(d) := k_0$. For $k \geq k_0 \geq 600$ and $n \geq k^{C_F(d)}$, we have $n \geq k^4 \geq \max\{16k, k^3\}$, so Lemma 2.4 applies with the sharper SUM budget (18). Both host sizes in Equation (11) are at least

$$\frac{1}{2}\sqrt{\frac{n}{k}} \geq \frac{1}{2}k^{c'_d+3/2} \geq k^{c'_d},$$

where the first inequality uses $C_F(d) \geq 2c'_d + 4$. Thus the conjecture applies at both hosts.

The projections of Theorem 2.3 and the transfer bound (19) give exponent $\gamma_d k/4$ in bases n for OV and n/k for XOR and SUM, with odd k for XOR. The range $n \geq k^4$ also gives $\log k \leq \frac{1}{4}\log n$, hence $\log(n/k) = \log n - \log k \geq \frac{3}{4}\log n$. Converting the XOR and SUM bounds to base n therefore gives all three bounds in the stated form. □

5.2 Open problems

Remaining depths and orientations. Prove the Π_3 and fixed-depth $d \geq 4$ cases of Conjecture 5.1, or obtain the corresponding target bounds directly. For Π_3, an analogue of the proof of Theorem 4.8 would need a distribution on maximal negative inputs and a bound for one middle DNF. **Even-parameter XOR.** At depth two, the supplementary bounds give an exponent linear in K for every sufficiently large even K, and a quadratic exponent when $K \equiv 2 \pmod 4$ (Corollary D.3 and Corollary D.2(iv)). Can one obtain a quadratic exponent for every sufficiently large even K at the principal row scale, and exponents linear in K at the remaining depths and orientations? A different lift with odd $r = \Omega(K)$ for every even K, or a direct projection enforcing exact cardinalities, would bypass the arithmetic obstruction of the current lift (Appendix D.2).

6 Acknowledgments

I am deeply grateful to my advisor, Professor Prashant Nalini Vasudevan, who encouraged me to study average-case planted k-XOR lower bounds in AC^0 and with whom I had many formative discussions throughout 2023. His guidance and support have been the foundation of this work. I also thank Professor Alexander Golovnev for a discussion in the summer of 2024 on possible routes to AC^0 lower bounds for the average-case planted problem, and for suggesting candidate source problems such as the Coin Problem; that exchange prompted the broader search that led to colored subgraph isomorphism.

7 AI Disclosure

Five specific technical contributions were aided by AI language models. We verified each of them independently before inclusion. (i) GPT-5.5 pointed us to Beame's small-clique lower bound for CRCW PRAMs [Bea90] as a candidate unconditional source statement for the growing-k, every-fixed-depth regime; we verified its applicability and developed the resulting argument (Lemma 4.4, Theorem 4.5) in Section 4.2. (ii) While we were attempting to extend our depth-two counting argument, GPT-5.6 Sol suggested combining two ideas for the depth-three result in Section 4.4: an adaptive process for revealing planted labels (Appendix F.1), which is related to the decision-tree querying in the algorithmic proofs of the switching lemma due to Razborov and to Beame [Raz95, Bea94],

and the formulation of the target as a minterm event on a random background, which draws on the planted-subgraph framework of [Ros08, LRR17]. We verified the suggestion independently and then developed the depth-three results for our padded expander family. (iii) While we were attempting to derive a direct projection for even k, GPT-5.5 suggested the odd-to-even lifting approach recorded as Lemma D.1 (Appendix D.1); we verified the construction and its arithmetic conditions before developing the consequences. (iv) GPT-5.5 suggested the constant-factor comparison and we derived it in the regular-core ceiling (Appendix B). (v) GPT-5.5 also suggested and derived the depth-sensitive exponent comparison at the principal bit-width for k-SUM in Appendix E.2, contrasting the depth-three block-carry construction against the direct depth-two verifier. In every case, the authors take full responsibility for the correctness of the final statements and proofs.

References

[ABDN18] Amir Abboud, Karl Bringmann, Holger Dell, and Jesper Nederlof. More consequences of falsifying SETH and the orthogonal vectors conjecture. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2018)*, pages 253–266. Association for Computing Machinery, 2018.

[Ajt83] Miklós Ajtai. Σ^1_1-formulae on finite structures. *Annals of Pure and Applied Logic*, 24(1):1–48, 1983.

[Alo21] Noga Alon. Explicit expanders of every degree and size. *Combinatorica*, 41(4):447–463, 2021.

[Ama10] Kazuyuki Amano. k-subgraph isomorphism on AC^0 circuits. *Computational Complexity*, 19(2):183–210, 2010. Preliminary version in CCC'09.

[ASS+24] Shweta Agrawal, Sagnik Saha, Nikolaj I. Schwartzbach, Akhil Vanukuri, and Prashant Nalini Vasudevan. k-SUM in the sparse regime: Complexity and applications. In Leonid Reyzin and Douglas Stebila, editors, *Advances in Cryptology—CRYPTO 2024, Part II*, volume 14921 of *Lecture Notes in Computer Science*, pages 315–351. Springer, 2024.

[Bea90] Paul Beame. Lower bounds for recognizing small cliques on CRCW PRAM's. *Discrete Applied Mathematics*, 29(1):3–20, 1990.

[Bea94] Paul Beame. A switching lemma primer. Technical Report UW-CSE-95-07-01, Department of Computer Science and Engineering, University of Washington, November 1994.

[BGGS16] Arnab Bhattacharyya, Ameet Gadekar, Suprovat Ghoshal, and Rishi Saket. On the hardness of learning sparse parities. In Piotr Sankowski and Christos Zaroliagis, editors, *24th Annual European Symposium on Algorithms (ESA 2016)*, volume 57 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 11:1–11:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016.

[CLSS26] Tameem Choudhury, Nutan Limaye, Karteek Sreenivasaiah, and Srikanth Srinivasan. New and improved concrete lower bounds for orthogonal vectors. Technical Report TR26-121, Electronic Colloquium on Computational Complexity, July 2026. Also available as arXiv:2607.23799.

[CS26] Tameem Choudhury and Karteek Sreenivasaiah. Depth-3 circuit lower bounds for k-OV. *ACM Transactions on Computation Theory*, 2026. Published online May 2026; preliminary version appeared in STACS 2024.

[CW19] Lijie Chen and R. Ryan Williams. An equivalence class for orthogonal vectors. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2019)*, pages 21–40. Society for Industrial and Applied Mathematics, 2019.

[DKK24] Itai Dinur, Nathan Keller, and Ohad Klein. Fine-grained cryptanalysis: Tight conditional bounds for dense k-SUM and k-XOR. *Journal of the ACM*, 71(3):23:1–23:41, 2024.

[DLSV25] Mina Dalirrooyfard, Andrea Lincoln, Barna Saha, and Virginia Vassilevska Williams. Average-case hardness of parity problems: Orthogonal vectors, k-SUM and more. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2025)*, pages 4613–4643. Society for Industrial and Applied Mathematics, 2025.

[EKMS22] Gregory Emdin, Alexander S. Kulikov, Ivan Mihajlin, and Nikita Slezkin. CNF encodings of parity. In Stefan Szeider, Robert Ganian, and Alexandra Silva, editors, *47th International Symposium on Mathematical Foundations of Computer Science (MFCS 2022)*, volume 241 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 47:1–47:12. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.

[FSS84] Merrick Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17:13–27, 1984.

[Hås86] Johan Håstad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing (STOC 1986)*, pages 6–20, 1986.

[Hås89] Johan Håstad. Almost optimal lower bounds for small depth circuits. In Silvio Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 143–170. JAI Press, 1989.

[IP01] Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-SAT. *Journal of Computer and System Sciences*, 62(2):367–375, 2001.

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, 2001.

[Kar72] Richard M. Karp. Reducibility among combinatorial problems. In Raymond E. Miller and James W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103, New York, 1972. Plenum Press.

[KW19] Daniel M. Kane and R. Ryan Williams. The orthogonal vectors conjecture for branching programs and formulas. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:15. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2019.

[LRR17] Yuan Li, Alexander Razborov, and Benjamin Rossman. On the AC^0 complexity of subgraph isomorphism. *SIAM Journal on Computing*, 46(3):936–971, 2017.

[MOP20] Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. Explicit near-Ramanujan graphs of every degree. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, pages 510–523, New York, NY, USA, 2020. Association for Computing Machinery. Corrected full version: arXiv:1909.06988v3.

[PPZ99] Ramamohan Paturi, Pavel Pudlák, and Francis Zane. Satisfiability coding lemma. *Chicago Journal of Theoretical Computer Science*, 1999(11):1–19, 1999.

[Raz95] Alexander A. Razborov. Bounded arithmetic and lower bounds in Boolean complexity. In Peter Clote and Jeffrey B. Remmel, editors, *Feasible Mathematics II*, volume 13 of *Progress in Computer Science and Applied Logic*, pages 344–386. Birkhäuser Boston, 1995.

[Ros08] Benjamin Rossman. On the constant-depth complexity of k-clique. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, STOC 2008, pages 721–730, New York, NY, USA, 2008. Association for Computing Machinery.

[Ros21] Gregory Rosenthal. Beating treewidth for average-case subgraph isomorphism. *Algorithmica*, 83(8):2521–2551, 2021.

[RW25] R. Ryan Williams. The orthogonal vectors conjecture and nonuniform circuit lower bounds. *SIAM Journal on Computing*, pages FOCS24–1–FOCS24–28, 2025. Special Section FOCS 2024; published online 14 October 2025.

[VW18] Virginia Vassilevska Williams. On some fine-grained questions in algorithms and complexity. In *Proceedings of the International Congress of Mathematicians (ICM 2018)*, volume 3, pages 3447–3487, 2018.

[Wag02] David Wagner. A generalized birthday problem. In Moti Yung, editor, *Advances in Cryptology—CRYPTO 2002*, volume 2442 of *Lecture Notes in Computer Science*, pages 288–304. Springer, 2002.

[Yao85] Andrew Chi-Chih Yao. Separating the polynomial-time hierarchy by oracles. In *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1985)*, pages 1–10, 1985.

A Source Functions and Cited Lower Bounds

We check the source and circuit-size conventions in the cited LRR and Beame bounds, then explain the host-size limitation of the published LRR switching estimate. As in Section 2.1, N is the source host size; n counts vectors per OV array or entries in an XOR or SUM list.

A.1 Coordinate equivalence with the LRR structured problem

The source in Definition 2.1 differs from LRR's colored subgraph problem only in its input-bit names.

Proposition A.1 (Coordinate-level equivalence)**.** *For every simple loopless graph P with a fixed endpoint order $f = (f_L, f_R)$ for each edge and every $N \geq 1$, the function P-SUB$_N$ of Definition 2.1 coincides, after a bijection of Boolean input coordinates, with* SUBGRAPH$_{\mathrm{col},N}(P)$ *of [LRR17, Sec. 5, p. 960].*

Proof. The structured host in [LRR17, Sec. 5, p. 960] is $V(P) \times [N]$ with first-coordinate coloring, as in [LRR17, Def. 2.8(v)] at vertex exponent $\alpha \equiv 1$. Its edge coordinates are

$$\{(a,u),(b,w)\} \quad \text{for } \{a,b\} \in E(P),\ u,w \in [N]. \tag{75}$$

Simplicity and the fixed endpoint orders make

$$\{(f_{\mathrm{L}},u),(f_{\mathrm{R}},w)\} \longmapsto X_{f,u,w} \tag{76}$$

a bijection. A color-respecting copy selects $(a,\phi(a))$ in every class. Different colors give distinct host vertices even when their indices agree. Its required edges are present exactly when Equation (2) holds; isolated colors may be assigned arbitrary indices. □

A.2 Applying the fixed-pattern LRR bound

We identify LRR's hard distribution on our source bits and account for any additional input and negation gates in their size measure.

Proof of Theorem 4.1. Fix P, d as in the theorem. Our assumption of no isolated vertices avoids an unused boundary case at $\alpha \equiv 1$. Connectedness is unnecessary: LRR's parameters apply to disconnected patterns [LRR17, Rem. 2.7].

By [LRR17, Cor. 4.2], there is an edge weighting $\beta : E(P) \to [0,2]$ with $(\mathbf{1},\beta) \in \theta_{\mathrm{col}}(P)$ such that

$$\kappa_{\mathbf{1},\beta}(P) = \kappa_{\mathrm{col}}(P) = \kappa(P). \tag{77}$$

These parameters and $G_{\mathbf{1},\beta}(N)$ are those of [LRR17, Defs. 2.8 and 2.12]. The displays of their Theorems 3.11–3.12 write $G_{\alpha,\beta}(P)$; the definition and proofs identify its argument as the host parameter N. At vertex exponent one, its host is

$$\{(a,u) : a \in V(P),\ 1 \le u \le \lfloor N^1 \rfloor\} = V(P) \times [N], \tag{78}$$

with exactly the coordinates of Equation (75). By Proposition A.1, a circuit computing our source is correct on every input from this distribution and therefore meets the correctness requirement of the cited average-case theorem.

Their type-I model uses unbounded-fan-in AND and OR gates with NOT gates [LRR17, Sec. 3.2, p. 948], so we apply their Theorem 3.12. Allowing depth $d+2$ covers conventions that count input or leaf-negation layers; the bound holds at every fixed depth. Their size measure counts gates [LRR17, Footnote 4, p. 940], without specifying whether inputs, constants or input negations count. For $S := \mathrm{size}_d(P\text{-SUB}_N)$ and $M_{\mathrm{src}} = e(P)N^2$ as in Section 2.2, our circuit uses at most $S + 2M_{\mathrm{src}} + 2$ gates under each convention: one variable and one shared negation node per source bit, and two constants.

If this additive quantity is at least $N^{\kappa(P)+1}$, the claim is immediate. Otherwise the converted circuit has polynomial size in N, as required by their type-I model. Theorem 3.12 and Equation (77) give

$$S + 2M_{\mathrm{src}} + 2 \ge N^{\kappa(P) - o_{P,d}(1)}. \tag{79}$$

For each fixed P and d, the exponent loss $o_{P,d}(1)$ tends to zero as $N \to \infty$, at a rate that may depend on P and d. Thus, for any $\eta > 0$, there is a threshold $N_0(P,d,\eta)$ beyond which this loss is at most η, proving Equation (37). For the final claim of Theorem 4.1, assume $\kappa(P) - \eta > 2$ and take N large enough that the exponent loss is at most $\eta/2$. Then

$$S \ge N^{\kappa(P)-\eta/2} - 2e(P)N^2 - 2 \ge N^{\kappa(P)-\eta},$$

since the convention term has exponent 2 and the final bound retains $\eta/2$ slack. □

Alternative proof. The equivalence between structured and unstructured inputs in [LRR17, Sec. 5, p. 960] also permits using their worst-case corollary. The proof above identifies the hard distribution directly on our input coordinates.

A.3 Beame's bound under the gate-count convention

Beame's bound gives a lower bound on the number of shared memory cells needed to detect an r-vertex clique on a parallel random-access machine (PRAM), even with arbitrarily many processors. We simulate a depth-d circuit on this machine in $O(d)$ parallel steps, using one memory cell per input or gate and one processor per wire. The memory lower bound then forces the circuit to have many gates; the number of wires affects only the processor count, which is unrestricted.

Proof of Lemma 4.4. Beame's PRIORITY CRCW PRAM allows concurrent reads and concurrent writes to shared memory. The priority rule means that, when several processors write to the same cell, the write by the least-index processor takes effect. The machine stores an N-vertex graph in $\binom{N}{2}$ Boolean input cells, initializes other cells to 0, and must write the answer to the designated input cell C_1. Processor transition functions are unrestricted: for each input length, we may build a machine whose rules encode the circuit for that length, without requiring a single algorithm to generate these rules.

In [Bea90, Sec. 2], $n = N$, $m = \binom{N}{2}$, $k = r$, and his $Clique_k^n$ is our K_r-CLIQUE$_N$; the phrase "n inputs" in Corollary 3.8 is a notational slip. We use Theorem 3.1(b) and its proof, which bound memory cells even with infinitely many processors, provided r is above an absolute constant and at most $\log N$. Corollary 3.8's circuit simulation does not specify gate versus wire size, so we give the simulation explicitly.

Let a depth-d circuit for K_r-CLIQUE$_N$ have G non-input gates. Propagate constants and flatten adjacent gates of the same type, without increasing the gate count or depth. Reserve G cells for gate values, one cell per input variable, and one auxiliary cell whose contents will be ignored. Use one processor per wire, together with one processor to copy the final answer.

We organize the gates into layers so that all children of a gate are evaluated before that gate is processed. Put input literals in layer 0, and assign each gate to one plus the largest layer among its children. Gates in the same layer can then be evaluated in parallel, since their inputs come from earlier layers. There are at most d layers of gates. Each wire processor is programmed with its destination gate's layer and keeps a local step counter, so it knows in advance when to read its child's value.

The simulation stores an OR gate's value and an AND gate's complement. For example, for an AND gate $g = h_1 \wedge h_2$, its cell stores $\neg g = \neg h_1 \vee \neg h_2$, which is 1 exactly when at least one child is false. Thus, for either gate type, a single child can certify that the stored bit should be 1: a true child certifies this for OR, and a false child certifies it for AND. The cell's initial 0 is correct if no child gives such a certificate.

Each PRAM step consists of a read, a local state update, and a write by every processor. At step t, call a wire processor active if its destination gate lies in layer t; only these processors test their children's values. Every predecessor gate lies in an earlier layer, so its cell already contains its correctly encoded value. In particular, a 0 read at this scheduled step is a valid encoded value even if the child cell has retained its initial 0; cells belonging to unevaluated gates are never used as inputs at this step.

An active processor writes 1 to its gate's cell exactly when its child's value supplies the certificate just described. For an internal wire, the alternating gate types make this equivalent to reading 0: an AND child's stored 0 means that child is true, as needed by an OR parent, while an OR child's

stored 0 means that child is false, as needed by an AND parent. For an input wire, the processor reads the underlying variable and accounts for any negation at that input leaf, testing whether the resulting literal is true for an OR parent or false for an AND parent.

Processors whose destination gates are outside the current layer, and active processors whose children supply no certificate, write 0 to the auxiliary cell. Hence every write to a gate cell is 1, so concurrent writes agree regardless of processor priority; if no processor writes to that cell, it retains its initial 0. After a gate's layer has been evaluated, the processors on its incoming wires send all subsequent writes to the auxiliary cell. With no further writes to the gate cell, it retains its computed value for use by later layers.

The output is evaluated in its own gate cell because C_1 must retain its input bit during the simulation. After at most d steps for the circuit layers, one further step copies the output to C_1, complementing the stored bit if the output is an AND gate. Thus the total number T of parallel steps satisfies $T \le d + 1 \le d + 2$, and the number of memory cells is

$$c(N) = G + \binom{N}{2} + O(1).$$

For $N \ge 3$, the input cells alone ensure $c(N) \ge N$, as required in Beame's argument.

To apply the memory lower bound, put $a := \log_N c(N)$, so that the simulation uses N^a cells in T parallel steps. Theorem 3.1(b) states $a \ge r/(43T^3)$, but its three terminal cases justify the weaker $a \ge r/(1000T^3)$ for every $T \ge 1$. Indeed, the first case gives $a \ge r/6$; the second gives $T \ge (3/10)(r/a)^{1/3} - 2$, hence

$$a \ge \frac{27r}{1000(T+2)^3} \ge \frac{r}{1000T^3};$$

and the final case gives $343T^3a/(8r) \ge 343/344$, hence $a \ge r/(43T^3)$. The second case reaches the published constant only for $T \ge 40$, so retaining 1000 avoids a small-depth gap. Consequently, with $b_d := 1/(1000(d+2)^3)$,

$$c(N) \ge N^{b_d r}.$$

We must now subtract the $\binom{N}{2}$ input cells to obtain a lower bound on G. Beame's argument requires both r and N to exceed absolute thresholds. Choose $r_\star$ large enough to exceed the threshold for r and to make $2^{r_\star}$ exceed the threshold for N. Whenever $r_\star \le r \le \log N$, we have $N \ge 2^r \ge 2^{r_\star}$, so both requirements hold. Set

$$r_d := \max\{r_\star, \lceil 3/b_d \rceil\}.$$

For every $r \ge r_d$, the exponent $b_d r$ is at least 3, so the memory lower bound dominates the $O(N^2)$ cells allocated to inputs and the auxiliary cell. We can therefore choose $N_d \ge 3$, depending only on d, large enough that $\binom{N}{2} + O(1) \le \frac{1}{2}N^{b_d r}$ for all $N \ge N_d$ and all $r \ge r_d$. To absorb the remaining factor $1/2$, set $a_d := b_d/2$. Increasing N_d if necessary also ensures $N^{(b_d - a_d)r} \ge 2$ throughout this range. Thus, for every $N \ge N_d$ and $r_d \le r \le \log N$,

$$G \ge N^{b_d r} - \binom{N}{2} - O(1) \ge \tfrac{1}{2}N^{b_d r} \ge N^{a_d r}.$$

This proves Lemma 4.4 with its stated constants and range; in particular, $a_d = \Theta(d^{-3})$ and $r_d = \Theta(d^3)$. Every application here fixes d. □

A.4 Host size in the cited switching estimate

The threshold in Theorem 4.1 may depend arbitrarily on the fixed pattern $P_k^{\exp}$. In particular, it gives no bound $N_0(P_k^{\exp}, d, \eta) \le k^{O_d(1)}$ of the kind needed in Conjecture 5.1. The calculation below explains why directly tracking the published switching estimate also fails at polynomial host size.

The relevant bad-event bound in [LRR17, Appendix C] has the form

$$B_{d,\delta}(S, N) := S\big(5N^{-\delta/d}\delta \log N\big)^{\delta \log N}. \tag{80}$$

Logarithms are base two as in Section 2.2; changing a fixed logarithm base changes only constants depending on d and δ in the estimates below. The fixed-pattern analysis takes $S = N^{O(1)}$ and chooses a constant $\delta > 0$ after fixing the pattern. With δ fixed independently of k, the next lemma shows that this estimate does not handle size $N^{\Theta(k)}$ at polynomial host size.

Lemma A.2 (Scale limitation of the published switching estimate)**.** *Fix constants $d \ge 1$, $\gamma > 0$, and $\delta > 0$. Along every sequence with $k \to \infty$ and $N \to \infty$,*

$$\log B_{d,\delta}(N^{\gamma k}, N) = \gamma k \log N - \Theta_{d,\delta}\big((\log N)^2\big) + O_\delta\big(\log N \log\log N\big). \tag{81}$$

Consequently, for every fixed $C > 0$,

$$B_{d,\delta}(N^{\gamma k}, N) \longrightarrow \infty \qquad \text{as } k \to \infty \text{ along } N = k^C. \tag{82}$$

Moreover, along every sequence with $k \to \infty$ and $N \to \infty$ for which $B_{d,\delta}(N^{\gamma k}, N) = o(1)$, one necessarily has

$$\log N = \Omega_{d,\delta,\gamma}(k). \tag{83}$$

Proof. Substituting $S = N^{\gamma k}$ into Equation (80) and taking logarithms gives

$$\begin{aligned}\log B_{d,\delta}(N^{\gamma k}, N) &= \gamma k \log N + \delta \log N \left(\log(5\delta) + \log\log N - \frac{\delta}{d}\log N\right) \\ &= \gamma k \log N - \frac{\delta^2}{d}(\log N)^2 + \delta \log N \log\log N + O_\delta(\log N),\end{aligned} \tag{84}$$

which proves Equation (81). If $N = k^C$, the first term in Equation (84) is $\Theta_{C,\gamma}(k \log k)$. The negative quadratic term is $O_{C,d,\delta}((\log k)^2)$, and the remaining terms are $O_{C,\delta}(\log k \log\log k)$. The logarithm therefore tends to $+\infty$, proving Equation (82).

Now consider a sequence with $k \to \infty$, $N \to \infty$, and $B_{d,\delta}(N^{\gamma k}, N) = o(1)$. Writing $x := \log N$, Equation (84) and $\log B_{d,\delta}(N^{\gamma k}, N) < 0$ imply, after rearranging and dropping the nonpositive term $-\delta x \log x$, that for all sufficiently large terms of the sequence,

$$\gamma k x \le \frac{\delta^2}{d}x^2 + O_\delta(x). \tag{85}$$

After dividing by $x > 0$, the right-hand side is $O_{d,\delta}(x)$ because $x \to \infty$. Thus $\gamma k \le K_{d,\delta} x$ for some constant $K_{d,\delta} > 0$ and all sufficiently large terms, giving $x = \Omega_{d,\delta,\gamma}(k)$ and proving Equation (83). □

Scope of the calculation. Lemma A.2 concerns only the published estimate with fixed δ; it does not rule out other proofs of Conjecture 5.1, including refinements of the method. At the hosts in Equation (11), the benchmark $k = \Theta(\log^2 n)$ gives $\log N = \Theta(\log n) = \Theta(\sqrt{k})$, below the requirement $\log N = \Omega_{d,\delta,\gamma}(k)$ in Equation (83). The auxiliary restriction parameter and constants in the probabilistic estimates of [LRR17, Appendix C] are also chosen after fixing the pattern and would need uniform bounds. The hitting-set estimate $|\mathcal{H}| \le 2^{|E(P)|}$ in the proof of LRR's type-II theorem [LRR17, Thm. 3.11] can, however, be absorbed into a source exponent $\gamma_d k$: when $e(P) = k - 1$ and $\log N \to \infty$, it contributes only $2^{O(k)} = N^{O(k/\log N)} = N^{o(k)}$.

B Pattern Constructions and Constants

We construct the two source pattern families used in the lower bounds, adding disjoint matching edges to reach exactly $k-1$ edges. The expander construction proves Lemma 2.2 and eq. (38); the clique family supplies the source for Theorem 4.5. We use the matching J_t and endpoint orders of Section 2.1.

B.1 The bounded-degree expander family

We build a bipartite 4-regular core with constant expansion, which gives a lower bound on its hardness parameter κ (Section 4.1.1). Adding matching edges preserves this bound because the core remains a minor of the padded pattern.

Expansion. For a graph P with at least two vertices, define

$$h(P) := \min_{\substack{\varnothing \neq S \subseteq V(P) \\ |S| \leq v(P)/2}} \frac{|E_P(S, V(P) \setminus S)|}{|S|}. \tag{86}$$

Equivalently, minimize the cut size divided by $\min\{|S|, |V(P) \setminus S|\}$ over nonempty proper S, as in [LRR17, Thm. 4.9].

We use the following properties of κ from [LRR17, Thm. 4.9, Prop. 4.3 and Thm. 5.1]; see also [Ros21, Thm. 3.8(ii),(iii) and Footnote 2].

Theorem B.1 (κ from expansion, treewidth, and minors)**.** *For every simple loopless graph P with at least three vertices, at least one edge, and no isolated vertices,*

$$\kappa(P) \geq \frac{v(P)h(P)}{3\Delta(P)}. \tag{87}$$

Furthermore, $\kappa(P) \leq \mathrm{tw}(P) + 1$, where $\mathrm{tw}(P)$ denotes the treewidth of P. Finally, κ is minor-monotone: if F is a minor of P, then $\kappa(F) \leq \kappa(P)$.

The assumptions on P avoid boundary cases in the cited expansion argument. For K_2, one has $\kappa(K_2) = 0$, whereas the right-hand side of Equation (87) is $2/3$. Excluding isolated vertices avoids undefined degree ratios. Disconnected graphs are allowed, although their expansion is zero. Our expander cores and their padded patterns satisfy these assumptions.

For a Δ-regular graph F, the discrete Cheeger inequality in the same normalization is [Ros21, Fact 2.4 and p. 2525]

$$h(F) \geq \frac{\Delta - \lambda_2(F)}{2}, \tag{88}$$

where $\lambda_2(F)$ is the second-largest adjacency eigenvalue.

To obtain a regular base graph with a prescribed number of vertices, we use Alon's near-Ramanujan construction [Alo21, Thm. 1.3] in the following form. The cited statement does not assert simplicity; we explain below why the construction used here provides it.

Theorem B.2 (Explicit simple near-Ramanujan graphs of every degree and size)**.** *For every fixed integer $\Delta \geq 3$ and every $\varepsilon > 0$, there is a finite threshold $n_{\mathrm{A}}(\Delta, \varepsilon)$ such that, for every $v \geq n_{\mathrm{A}}(\Delta, \varepsilon)$ with $v\Delta$ even, there is a simple Δ-regular graph on v vertices whose nontrivial adjacency eigenvalues, that is, all adjacency eigenvalues other than one copy of Δ, have absolute value at most $2\sqrt{\Delta - 1} + \varepsilon$.*

Alon's construction starts from a near-Ramanujan graph on $v + o(v)$ vertices in which every neighborhood of a fixed radius, chosen in terms of Δ and ε, contains at most one cycle [Alo21, Thm. 3.3]. We take this starting graph from the corrected construction [MOP20, arXiv v3, Sec. 5], which supplies these spectral and local-cycle properties together with simplicity. Alon's procedure then deletes a small set of widely separated vertices to leave exactly v vertices [Alo21, Sec. 3].

Each deleted vertex has a cycle-free neighborhood, so no two of its immediate neighbors are adjacent: such an edge would form a triangle. Distinct deleted vertices are far enough apart that their sets of immediate neighbors are disjoint and no edge joins these sets [Alo21, Lem. 3.1(2),(3)]. Hence the immediate neighbors of all deleted vertices together form an independent set.

Each of these neighbors loses exactly one edge when the selected vertices are deleted. The parity condition $v\Delta$ even ensures that the number of affected neighbors is even, so Alon's procedure pairs them up and adds one edge within each pair. Every affected vertex then regains its missing edge, restoring degree Δ. Each pair consists of distinct vertices that were previously nonadjacent, so the added edges create neither loops nor parallel edges. This verifies simplicity for the graph to which Alon's spectral bound applies.

The finite threshold. We use only the existence of the threshold $n_{\mathrm{A}}(\Delta, \varepsilon)$ in Theorem B.2, not the construction's efficiency: our nonuniform projections fix the pattern before reading the source input. We do not optimize this threshold.

Proof of Lemma 2.2 and Equation (38). Fix the constant spectral bound $\Lambda := 2\sqrt{3}+1/10 < 4$ and set $n_* := \max\{n_{\mathrm{A}}(4, 1/10), 5\}$. The threshold k_0 of Section 2.1 may be chosen at least $\max\{4n_* + 1, 600\}$. For $k \geq k_0$, put

$$r_k := \left\lfloor \frac{k-1}{4} \right\rfloor, \qquad t_k := k - 1 - 4r_k \in \{0, 1, 2, 3\}. \tag{89}$$

By Theorem B.2, there is a simple 4-regular graph Q_k on r_k vertices with every nontrivial eigenvalue bounded in absolute value by Λ. The bound $\Lambda < 4$ forces Q_k to be connected: each connected component of a 4-regular graph contributes an independent eigenvector with eigenvalue 4, namely its indicator vector, whereas only one eigenvalue is allowed to equal 4. It also rules out bipartiteness, since the vector equal to 1 on one bipartition side and -1 on the other would have eigenvalue -4.

To construct a bipartite core, take two copies of $V(Q_k)$, forming the sides $I_k := V(Q_k) \times \{0\}$ and $I'_k := V(Q_k) \times \{1\}$. For every edge $\{a, b\}$ of Q_k, add the two edges $\{(a, 0), (b, 1)\}$ and $\{(b, 0), (a, 1)\}$. Call the resulting graph $F_k^{\exp}$; this construction is the *bipartite double cover* of Q_k. Each vertex copy has one neighbor for each neighbor of the original vertex, so the graph is 4-regular. All edges run between the two sides, and the absence of repeated edges in Q_k ensures that the new graph is simple. If A_k is the adjacency matrix of Q_k, the adjacency matrix of $F_k^{\exp}$ is

$$\begin{pmatrix} 0 & A_k \\ A_k & 0 \end{pmatrix}. \tag{90}$$

If $A_k z = \lambda z$, then (z, z) is an eigenvector of this block matrix with eigenvalue λ, while $(z, -z)$ has eigenvalue $-\lambda$. Applying this observation to an eigenbasis of A_k gives all $2r_k$ eigenvalues of $F_k^{\exp}$: there is one eigenvalue 4, one eigenvalue -4, and every remaining eigenvalue has absolute value at most Λ. In particular, the eigenvalue 4 has multiplicity one. Since every connected component would contribute an independent eigenvector with eigenvalue 4, the core is connected. Its second-largest eigenvalue satisfies $\lambda_2(F_k^{\exp}) \leq \Lambda$, while -4 is its smallest eigenvalue.

Add the t_k disjoint matching edges needed to reach exactly $k-1$ edges, defining

$$P_k^{\exp} := F_k^{\exp} \dot\cup J_{t_k}. \tag{91}$$

The padded pattern has no isolated vertices, and

$$v(F_k^{\exp}) = 2r_k, \quad e(F_k^{\exp}) = 4r_k > k-5, \quad |I_k| = |I_k'| = r_k > \frac{k-1}{4} - 1. \tag{92}$$

Thus $e(P_k^{\exp}) = k-1$ and $\Delta(P_k^{\exp}) = 4$. By Equation (88), the core has expansion at least $(4-\Lambda)/2 > 0$. The core remains a minor after padding, so Theorem B.1 gives

$$\kappa(P_k^{\exp}) \geq \kappa(F_k^{\exp}) \geq \frac{2r_k(4-\Lambda)}{24} > \frac{r_k}{30}, \tag{93}$$

where $\sqrt{3} < 7/4$ gives $4-\Lambda > 2/5$. Rounding down in the definition of r_k loses at most 1. Since $k \geq 600$ gives $1 \leq (k-1)/64$, we obtain

$$\kappa(P_k^{\exp}) > \frac{r_k}{30} \geq \frac{1}{30}\left(\frac{k-1}{4} - \frac{k-1}{64}\right) = \frac{k-1}{128}. \tag{94}$$

This proves Equation (38) with $c_\kappa = 1/128$; $k_0 \geq 600$ also gives $c_\kappa(k_0-1) > 4$.

Finally, to transfer circuit lower bounds from the core problem to the padded problem, keep every core input variable and set all padding-edge variables to 1. This gives the projection

$$F_k^{\exp}\text{-SUB}_N \leq_{\text{proj}} P_k^{\exp}\text{-SUB}_N \qquad (N \geq 1). \tag{95}$$

To verify correctness, suppose the core instance contains a colored copy, witnessed by a map $\phi : V(F_k^{\exp}) \to [N]$. Extend ϕ by assigning index 1 to each new matching color. Every required matching edge is present because its variables were fixed to 1, so the extended map witnesses a colored copy of the padded pattern. Conversely, any colored copy of the padded pattern restricts to a colored copy of the core, whose variables were left unchanged. Thus the original core instance accepts exactly when the constructed padded instance accepts. Keeping the core endpoint orders preserves every core input coordinate. □

Treewidth. For an expander core $F = F_k^{\exp}$, Equation (93) and Theorem B.1 give $\kappa(F) = \Omega(v(F))$ and $\kappa(F) \leq \text{tw}(F) + 1 \leq v(F)$. Thus $\kappa(F) = \Theta(\text{tw}(F)) = \Theta(v(F))$; Theorem 4.1 uses κ.

A limit of regular-core patterns. The same treewidth bound limits the exponent that the LRR theorem can yield from a regular core padded by a matching. For $P = F \dot\cup J_t$ with nonempty Δ-regular F, $\Delta \geq 2$, and $e(P) = k-1$, Theorem B.1 gives

$$\kappa(P) \leq \text{tw}(P) + 1 \leq v(F) = 2e(F)/\Delta \leq 2(k-1)/\Delta. \tag{96}$$

Indeed, F contains a cycle, so $\text{tw}(F) \geq 2$, whereas $\text{tw}(J_t) \leq 1$. Treewidth is the maximum over components, and $\text{tw}(F) \leq v(F) - 1$. Even a source bound $N^{\kappa(P)}$ would therefore give a leading target exponent of at most $(k-1)/\Delta$ after the quadratic host substitution (Lemma 2.4), for fixed k as $n \to \infty$. At the principal budgets, witness enumeration in Proposition E.1 has leading exponent k in bases n for OV and n/k for XOR. The ratio of this enumeration exponent to the ceiling $(k-1)/\Delta$ is $\Delta k/(k-1) > \Delta$, hence greater than 4 for our core. This limitation concerns only the LRR theorem's κ-based bounds for regular cores padded by matchings.

B.2 The clique family

We transfer ordinary clique hardness to colored subgraph isomorphism, then add matching edges to meet the projection requirement $e(P) = k - 1$ in Theorem 2.3.

Definition B.3 (Ordinary clique and padded clique pattern)**.** For integers $r \geq 2$ and $N \geq 1$, let K_r-CLIQUE$_N$ be the total monotone Boolean function on variables $y_{\{u,w\}}$, indexed by unordered pairs of distinct $u, w \in [N]$, that accepts exactly when the represented graph contains an r-clique.

For integers $k \geq 2$ and $r \geq 2$ with $\binom{r}{2} \leq k - 1$, define

$$P_{k,r}^{\mathrm{clq}} := K_r \,\dot\cup\, J_{k-1-\binom{r}{2}}. \tag{97}$$

Then $e(P_{k,r}^{\mathrm{clq}}) = k - 1$ and $\Delta(P_{k,r}^{\mathrm{clq}}) = r - 1$.

Lemma B.4 (Projections from ordinary clique to the padded pattern)**.** *For every k, r, N satisfying Definition B.3,*

$$K_r\text{-CLIQUE}_N \leq_{\mathrm{proj}} K_r\text{-SUB}_N \leq_{\mathrm{proj}} P_{k,r}^{\mathrm{clq}}\text{-SUB}_N. \tag{98}$$

Proof. The first projection makes one copy of the input graph's vertex set for each of the r clique colors. We copy the input adjacency between color classes and forbid a colored witness from using the same original vertex twice. The second projection adds the matching colors with every matching-edge variable fixed to 1, so these added colors can always be assigned host vertices.

For the first projection, let y encode an ordinary graph on $[N]$. For each edge f of the colored-clique pattern, using its fixed endpoint order, substitute

$$X_{f,u,w} \longmapsto \begin{cases} y_{\{u,w\}}, & u \neq w, \\ 0, & u = w. \end{cases} \tag{99}$$

The zero entries for $u = w$ prevent two colors from selecting copies of the same input vertex. Thus an accepting colored copy chooses r distinct labels in $[N]$. Every pair of chosen labels is adjacent in the input graph, since its colored edge has value $y_{\{u,w\}} = 1$. These labels therefore form an ordinary r-clique. Conversely, given an ordinary r-clique, assign its vertices one to each color. Every required colored edge is present by the substitution, so the colored instance accepts. Each substituted coordinate is an input bit or a constant, as required for a depth-zero projection.

For the second projection, start from a colored-clique instance, keep its edge variables, and set every variable for an added matching edge to 1. A colored clique extends to a copy of the padded pattern by assigning host index 1 to every new matching color, because all required matching edges are present. Conversely, any colored copy of the padded pattern includes a colored clique on its original clique colors. The padded instance therefore accepts exactly when the colored-clique instance accepts, proving the second projection. □

C Lower Bounds for Other Encoding Budgets

Lemma 2.4 uses the target parameters n, k to choose a source host, then bounds the encoding budget needed to represent it. Here we reverse that calculation: for a given OV dimension D, XOR row count m, or SUM bit width m, we choose a source host that fits the available budget. This yields a target lower bound when the host is also large enough for a source lower-bound theorem. We first derive feasible hosts, then check the thresholds for fixed and growing patterns.

C.1 Choosing the source host

Larger source hosts give stronger lower bounds, but require more target entries and more encoding coordinates. The host size is therefore limited both by the capacity of the arrays or list and by the available D or m in Theorem 2.3. We make these constraints explicit and choose an integer N satisfying both. The extension arguments then reach the prescribed D or m whenever it exceeds the chosen host's encoding requirements.

For $k \geq 2$ and $e(P) = k - 1$, Equations (21), (27) and (31), together with $|\mathcal{I}(P)| \leq 2(k-1)$ from Equation (10), $R(P) \leq 3k$ from Equation (30), and $\ell_N \leq \log(2N)$, give

$$m_{\oplus}(P, N) \leq 4(k-1)\log(2N) + 2(k-1) + 3 \leq 8k\log(2N), \tag{100}$$

$$m_{\Sigma}(P, N) = R(P)\lceil \log(2kN+1) \rceil \leq 9k\log(2kN), \tag{101}$$

$$D_0(P, N, n) \leq 4k\ell_N + k + 2\ell_n. \tag{102}$$

The first bound includes all XOR rows; the second uses $\lceil \log(2kN+1) \rceil \leq 3\log(2kN)$. To measure the available $\log N$, write $[x]_+ := \max\{x, 0\}$ and define the width parameters

$$w_{\oplus} := \min\{\log(n/k), m/k\}, \tag{103}$$

$$w_{\Sigma} := \left[\min\{\log(n/k), m/(9k) - \log(2k)\}\right]_+, \tag{104}$$

$$w_{\mathrm{OV}} := \left[\min\{\tfrac{1}{2}\log n, (D - k - 2\ell_n)/(4k)\}\right]_+. \tag{105}$$

Feasible hosts. The capacity and encoding constraints place upper bounds on N. We choose N to equal the expression in the middle column of Table 2; the last column gives a guaranteed lower bound for this choice.

Target	Chosen host	Guaranteed lower bound
Odd k-XOR	$N = \left\lfloor \min\{2^{m/(8k)-1}, \sqrt{(n-1)/(k-1)}\} \right\rfloor$	$\log N \geq w_{\oplus}/16$
k-SUM	$N = \left\lfloor \min\{2^{m/(9k)}/(2k), \sqrt{(n-1)/(k-1)}\} \right\rfloor$	$\log N \geq w_{\Sigma}/4$
k-OV	$N = 2^{\lfloor w_{\mathrm{OV}} \rfloor}$	$\log N \geq w_{\mathrm{OV}} - 1$

Table 2: Chosen hosts within the target's capacity and encoding budget, with lower bounds on their logarithms. The XOR, SUM, and OV rows are valid when $w_{\oplus} \geq 32$, $w_{\Sigma} \geq 4$, and $w_{\mathrm{OV}} > 0$, respectively.

For both list targets, the square-root term enforces the capacity constraint: the projection produces $1 + (k-1)N^2$ non-padding objects, which must fit into a list of length n. Choosing $N \leq \sqrt{(n-1)/(k-1)}$ ensures precisely this inequality.

The other term in each minimum enforces the encoding budget. For XOR, it ensures $\log(2N) \leq m/(8k)$, so Equation (100) gives $m_{\oplus}(P, N) \leq m$. For SUM, it ensures $\log(2kN) \leq m/(9k)$, so Equation (101) gives $m_{\Sigma}(P, N) \leq m$. Thus the chosen list hosts meet both requirements of Theorem 2.3.

To compare the chosen hosts with the source theorems' thresholds, we need the lower bounds in the last column. Let $Y_{\oplus}$ and Y_{Σ} denote the respective minima before taking the floor. The XOR condition $w_{\oplus} \geq 32$ and the SUM condition $w_{\Sigma} \geq 4$ each imply $n \geq k$ by Equations (103) and (104). Hence $(n-1)/(k-1) \geq n/k$, so the common capacity term $\sqrt{(n-1)/(k-1)}$ has logarithm at least $\frac{1}{2}\log(n/k)$. The encoding limit has logarithm $m/(8k) - 1$ for XOR and $m/(9k) - \log(2k)$ for

SUM. The width definitions therefore give the following bounds on the logarithms of $Y_\oplus$ and Y_Σ:

$$\log Y_\oplus \geq \min\{w_\oplus/8 - 1, w_\oplus/2\} = w_\oplus/8 - 1 \geq 3,$$
$$\log Y_\Sigma \geq \min\{w_\Sigma, w_\Sigma/2\} = w_\Sigma/2 \geq 2.$$

Taking the floor only makes a small difference: for $Y \geq 2$, one has $\lfloor Y \rfloor \geq Y/2$, so rounding costs at most one unit in the base-two logarithm. Hence the chosen list hosts satisfy the lower bounds in Table 2:

$$\log\lfloor Y_\oplus \rfloor \geq w_\oplus/8 - 2 \geq w_\oplus/16,$$
$$\log\lfloor Y_\Sigma \rfloor \geq w_\Sigma/2 - 1 \geq w_\Sigma/4.$$

For OV, we check the two requirements of Theorem 2.3(i): each edge array must have room for its N^2 edge vectors, and the full vector encoding must fit within D coordinates. Choosing a power of two makes the label length exact: $\ell_N = \log N = \lfloor w_{\mathrm{OV}} \rfloor$. The two upper limits in Equation (105), together with the dimension bound Equation (102), verify these capacity and dimension requirements:

$$N^2 \leq n, \qquad D_0(P, N, n) \leq 4k\lfloor w_{\mathrm{OV}} \rfloor + k + 2\ell_n \leq D.$$

Also, rounding w_{OV} down loses at most one, so $\log N \geq w_{\mathrm{OV}} - 1$. Each chosen host therefore permits Theorem 2.3, preserving depth and orientation. The next subsection checks when these hosts are large enough for the source lower bounds.

C.2 Host-size requirements for fixed and growing patterns

A feasible host yields a target lower bound only once it reaches the source theorem's threshold. The lower bounds on $\log N$ in Table 2 then determine the target exponent in terms of the available budget. For fixed d and k, the threshold in Corollary 4.2 may depend on both parameters; the first two results below translate it into thresholds on $w_\oplus$ and w_Σ from Equations (103) and (104). For growing k, the source bounds in Theorems 4.6 and 4.8 instead require a host size at least a specified power of k. We translate those requirements into conditions on all three width parameters.

Theorem C.1 (General row count for odd k-XOR)**.** *There is a universal $c_\oplus > 0$ such that, for every fixed depth d and fixed odd $k \geq k_0$, a threshold $L_\oplus(d, k)$ satisfies*

$$w_\oplus \geq L_\oplus(d, k) \quad \Longrightarrow \quad \log \mathrm{size}_d(k\text{-}\mathsf{XOR}_{n,m,k}) \geq c_\oplus \min\{k \log(n/k), m\}. \tag{106}$$

Theorem C.2 (General bit width for k-SUM)**.** *There is a universal $c_\Sigma > 0$ such that, for every fixed depth d and fixed $k \geq k_0$ of either parity, a threshold $L_\Sigma(d, k)$ satisfies*

$$w_\Sigma \geq L_\Sigma(d, k) \quad \Longrightarrow \quad \log \mathrm{size}_d(k\text{-}\mathsf{SUM}_{n,m,k}) \geq c_\Sigma k w_\Sigma. \tag{107}$$

Proof of both theorems. Use the common pattern $P_k^{\exp}$ and fix one universal constant $0 < \gamma_\star < c_\kappa$, where c_κ is given in Equation (38). For fixed (d, k), let $N_0 = N_0(\gamma_\star, d, k) \geq 2$ be a threshold from Corollary 4.2. It suffices to choose

$$L_\oplus(d, k) := \max\{32, 16 \log N_0\}, \qquad L_\Sigma(d, k) := \max\{4, 4 \log N_0\}.$$

The constants 32 and 4 ensure feasibility, while the terms involving $\log N_0$ ensure that the chosen host in Table 2 satisfies $N \geq N_0$.

For either target, Theorem 2.3 and projection monotonicity transfer the source bound Equation (39), giving a lower bound $\gamma_\star(k-1)\log N$ on the logarithm of circuit size. Using $k-1 \geq k/2$ and substituting the respective lower bounds on $\log N$ from Table 2 gives

$$\begin{aligned}\log \mathrm{size}_d(k\text{-}\mathsf{XOR}_{n,m,k}) &\geq \frac{\gamma_\star}{32} k w_\oplus = \frac{\gamma_\star}{32}\min\{k\log(n/k), m\},\\ \log \mathrm{size}_d(k\text{-}\mathsf{SUM}_{n,m,k}) &\geq \frac{\gamma_\star}{8} k w_\Sigma.\end{aligned}$$

Thus $c_\oplus = \gamma_\star/32$ and $c_\Sigma = \gamma_\star/8$ are universal; only the thresholds $L_\oplus(d,k)$ and $L_\Sigma(d,k)$ depend on (d,k). □

Growing patterns. The growing-pattern source bounds require a host size at least a specified power of k. We again use $P_k^{\exp}$, with odd k required for the direct XOR statements. At depth two, Theorem 4.6(i) gives $\mathrm{size}_2(P_k^{\exp}\text{-SUB}_N) \geq N^{\gamma_D k}$ whenever $k \geq k_D$ and $N \geq k^2$. By Table 2, it suffices to require $w_\oplus/16 \geq 2\log k$, $w_\Sigma/4 \geq 2\log k$, or $w_{\mathrm{OV}} - 1 \geq 2\log k$, according to the target. These requirements also ensure feasibility. In the OV case they imply $w_{\mathrm{OV}} \geq 2$, so $\log N \geq w_{\mathrm{OV}} - 1 \geq w_{\mathrm{OV}}/2$. Transferring the source bound therefore gives, for $k \geq k_D$,

$$\begin{aligned}\log \mathrm{size}_2(k\text{-}\mathsf{XOR}_{n,m,k}) &\geq \frac{\gamma_D}{16} k w_\oplus \quad \text{if } w_\oplus \geq 32\log k,\\ \log \mathrm{size}_2(k\text{-}\mathsf{SUM}_{n,m,k}) &\geq \frac{\gamma_D}{4} k w_\Sigma \quad \text{if } w_\Sigma \geq 8\log k,\\ \log \mathrm{size}_2(k\text{-}\mathsf{OV}_{n,D,k}) &\geq \frac{\gamma_D}{2} k w_{\mathrm{OV}} \quad \text{if } w_{\mathrm{OV}} \geq 2\log k + 1.\end{aligned} \tag{108}$$

For the Σ_3 orientation, fix $0 < \gamma < 1/8$. Take $k_E(\gamma)$ and $C_E(\gamma)$ from Theorem 4.8, enlarging $C_E(\gamma)$ to exceed 2 if necessary. The source bound Equation (66) gives $\mathrm{size}_{\Sigma_3}(P_k^{\exp}\text{-SUB}_N) \geq N^{\gamma k}$ when $k \geq k_E(\gamma)$ and $N \geq k^{C_E(\gamma)}$. Thus we replace $2\log k$ in the depth-two host requirement by $C_E(\gamma)\log k$. The choice $C_E(\gamma) > 2$ ensures that, as at depth two, the resulting conditions imply feasibility and, for OV, $w_{\mathrm{OV}} \geq 2$. Applying Table 2 gives, for $k \geq k_E(\gamma)$,

$$\begin{aligned}\log \mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{XOR}_{n,m,k}) &\geq \frac{\gamma}{16} k w_\oplus \quad \text{if } w_\oplus \geq 16 C_E(\gamma)\log k,\\ \log \mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{SUM}_{n,m,k}) &\geq \frac{\gamma}{4} k w_\Sigma \quad \text{if } w_\Sigma \geq 4 C_E(\gamma)\log k,\\ \log \mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{OV}_{n,D,k}) &\geq \frac{\gamma}{2} k w_{\mathrm{OV}} \quad \text{if } w_{\mathrm{OV}} \geq C_E(\gamma)\log k + 1.\end{aligned} \tag{109}$$

For an even XOR target parameter K, the lift in Lemma D.1 applies when its arithmetic and list-length hypotheses hold, adding $O(K\log n')$ tag rows for target list length n'. The even-parameter consequences in Corollary D.2 account for this cost by requiring $m' \geq CK\log n'$ for a sufficiently large universal C; we do not optimize the tag cost for smaller budgets.

Larger budgets. The width parameters simplify once the encoding costs independent of N consume at most half the available budget. For SUM, the sufficient bit width in Equation (101) is $9k\log(2kN) = 9k\log(2k) + 9k\log N$. The first term explains the subtraction of $\log(2k)$ in Equation (104). If $m \geq 18k\log(2k)$, then

$$\frac{m}{9k} - \log(2k) \geq \frac{m}{18k}, \qquad w_\Sigma \geq \frac{1}{18}\min\{\log(n/k), m/k\}.$$

For fixed (d,k) with $k \geq k_0$, assume in addition that $\min\{\log(n/k), m/k\} \geq 18L_\Sigma(d,k)$. Then $w_\Sigma \geq L_\Sigma(d,k)$, and Theorem C.2 gives the simpler bound

$$\log \mathrm{size}_d(k\text{-}\mathsf{SUM}_{n,m,k}) \geq \frac{c_\Sigma}{18} \min\{k \log(n/k), m\}.$$

For OV, the cost $k + 2\ell_n$ in Equation (21) is independent of N: k coordinates enforce edge presence and exclude padding, and $2\ell_n$ enforce distinctness within each array. If $D \geq 2(k + 2\ell_n)$, then at least $D/2$ coordinates remain for comparing endpoint labels. Consequently, Equation (105) gives

$$k w_{\mathrm{OV}} \geq \min\{\tfrac{1}{2} k \log n, D/8\}.$$

For both OV bounds below, assume $D \geq 2(k + 2\ell_n)$. For depth two, assume also $k \geq k_D$ and the OV width condition in Equation (108); for Σ_3, assume $k \geq k_E(\gamma)$ and the OV width condition in Equation (109). Under these respective assumptions, the corresponding bounds simplify to

$$\log \mathrm{size}_2(k\text{-}\mathsf{OV}_{n,D,k}) \geq \frac{\gamma_D}{16} \min\{k \log n, D\},$$
$$\log \mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{OV}_{n,D,k}) \geq \frac{\gamma}{16} \min\{k \log n, D\}.$$

At fixed k, the cost $k + 2\ell_n$ is already $\Theta(\log n)$, the same order as the principal OV budget in Equation (15). Thus varying N in this construction does not give a smaller asymptotic dimension scale.

D Supplementary Results for k-XOR

We transfer odd-parameter XOR bounds to even parameters by repeating selections. We then give a supplementary PARITY projection using $O(k \log(n/k))$ rows and derive its depth-two consequence. A classical bound on isolated satisfying assignments limits the number of PARITY variables that any such projection can encode.

D.1 The even-K lifting projection

The lift replaces each source index by a group of t positions and adds s auxiliary positions. Extra rows force target witnesses to select whole groups, with the auxiliaries selected exactly when the number of selected groups is odd. These selection rules are linear equations, and a short probabilistic argument compresses them into $O(K \log n')$ fixed tag rows for target list length n'.[3]

Lemma D.1 (Odd-to-even lifting projection)**.** *Let $r \geq 1$ be odd, let $t \geq 2$ and $s \geq 0$ be integers, put $K := tr + s$, and assume that K is even and that either $s = 0$ or $t \nmid s$. For all $n, m \geq 1$ and every target list length $n' \geq tn + s$, there is an integer $q = O(K \log n')$ and a projection*

$$k\text{-}\mathsf{XOR}_{n,m,r} \leq_{\mathrm{proj}} k\text{-}\mathsf{XOR}_{n',m+q,K}. \tag{110}$$

Proof. Replace each source position i by a group of t target positions, putting a_i at the first position and 0^m at the others. The s auxiliary positions and the remaining padding also have data 0^m.

To enforce the selection rules, equate the binary selection indicators within each group, equate each auxiliary indicator with the XOR of one indicator per group, and set padding indicators to zero. These are homogeneous linear equations.

[3]The AI language model GPT-5.5 suggested this odd-to-even lifting approach, using repetition with auxiliary parity copies and a separating tag matrix; the authors independently verified the construction.

Use the coefficient vectors of q independent uniform $\mathbb{F}_2$-linear combinations of these equations as tag rows. A selection obeying the rules always passes; any fixed selection violating them passes all q rows with probability 2^{-q}. There are at most $(n')^K$ selections of size K, so taking the least integer $q > K \log n'$ makes the union bound less than one. Fix successful tag rows before reading the source input.

A surviving K-selection uses w whole groups and has size $tw + s(w \bmod 2)$. An odd w forces $w = r$. An even w would give $tw = tr + s$, which is impossible by $t \nmid s$ when $s > 0$ and by oddness of r when $s = 0$. Thus every target witness selects exactly r groups, with data XOR equal to the corresponding source XOR. Conversely, every source witness selects r whole groups and all auxiliaries, giving a target witness. Every output bit is a constant or a source bit, so this is a projection. Appending zero rows reaches every larger row budget. □

D.2 Lower bounds for even parameters

The lift turns a circuit for the even target parameter K into one for the odd source parameter r, without increasing size or depth or changing the top-gate type (projection monotonicity). To apply an odd-parameter lower bound, the target list must hold the groups of t positions and the s auxiliary positions, and the target rows must hold both the source rows and the tag rows. Once these requirements are met, the transferred exponent is governed by r.

Corollary D.2 (Lower bounds from an odd parameter)**.** *Let r, t, s, K satisfy the hypotheses of Lemma D.1. There is a universal constant $C > 0$ such that, for $n' \geq 4K$ and $m' \geq CK \log n'$, the following hold.*

(i) For fixed d, t, r, s with $r \geq k_0$, once n' is sufficiently large in terms of these parameters,

$$\mathrm{size}_d(k\text{-}\mathsf{XOR}_{n',m',K}) \geq (n'/K)^{\beta(r-1)/2}, \tag{111}$$

where β is the universal rate in Theorem 4.3.

(ii) If r is sufficiently large and n' is at least a sufficiently large universal power of K,

$$\mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{XOR}_{n',m',K}) \geq (n'/K)^{\Omega(r)}. \tag{112}$$

(iii) Assuming Conjecture 5.1, for fixed d, if r is sufficiently large in terms of d and n' is at least a sufficiently large power of K depending only on d,

$$\mathrm{size}_d(k\text{-}\mathsf{XOR}_{n',m',K}) \geq \exp(\Omega_d(r \log n')). \tag{113}$$

(iv) If $r \geq 3$ and $n' \geq 8K$, then

$$\mathrm{size}_2(k\text{-}\mathsf{XOR}_{n',m',K}) \geq 2^{\Omega(r^2 \log(n'/K))}. \tag{114}$$

Proof. Choose $n := \lfloor (n'-s)/t \rfloor$, the largest source list length for which the groups of t positions and the s auxiliary positions fit in the target. Thus $tn + s \leq n'$, as required by Lemma D.1. Flooring loses less than one, and $s + t \leq K \leq n'/4$, so $n > (n' - s - t)/t \geq n'/(2t)$. Since $tr \leq K$ and $n'/K \geq 4$, this gives

$$\frac{n}{r} \geq \frac{n'}{2tr} \geq \frac{n'}{2K} \geq \left(\frac{n'}{K}\right)^{1/2}. \tag{115}$$

This comparison of bases accounts for the factor $1/2$ in the exponent of (i).

Next reserve rows for both the source instance and the tags. Take $m := \lceil C_{\oplus} r \log(en/r) \rceil$, which meets the source row requirement in Theorems 4.3 and 4.8 and Corollary 5.2. Since $C_{\oplus} \geq 1$, it also meets the requirement $m \geq r \log(n/r)$ in Corollary D.5. The bounds $r \leq K$ and $n \leq n'$ give $m = O(K \log n')$, with a universal constant. The tag count q in Lemma D.1 has the same bound. A sufficiently large universal C therefore ensures $m + q \leq m'$. Apply the lift and append zero rows to reach exactly m' rows.

For (i), the lifting parameters are fixed, so n eventually reaches the threshold $n_0(d, r)$ in Theorem 4.3 as n' grows. The odd-parameter bound (41) gives $(n/r)^{\beta(r-1)}$. Transfer this through the lift and use Equation (115) to obtain the stated even-parameter bound.

For (ii), fix a positive constant rate allowed by Theorem 4.8. Since $n \geq n'/(2K)$ and $r \leq K$, taking n' to be a sufficiently large universal power of K meets that theorem's polynomial source-length threshold once r is sufficiently large. Its lower bound $(n/r)^{\Omega(r)}$ transfers through the lift with the Σ_3 orientation preserved, and Equation (115) gives (ii).

For (iii), assume Conjecture 5.1 and fix d. The same comparison meets the polynomial source-length threshold in Corollary 5.2 when n' is a sufficiently large power of K depending only on d. Increasing this power if necessary also ensures $n \geq \sqrt{n'}$. For r sufficiently large in terms of d, Equation (72) gives $\exp(\Omega_d(r \log n))$; transferring through the lift and using $\log n \geq \frac{1}{2} \log n'$ proves (iii).

For (iv), the stronger length condition $n' \geq 8K$ gives $n/r \geq n'/(2K) \geq 4$. Thus Corollary D.5 applies to the odd source parameter $r \geq 3$, giving a depth-two lower bound $2^{\Omega(r^2 \log(n/r))}$. Transfer it through the lift and use Equation (115) to obtain (iv). □

For $K \equiv 2 \pmod 4$, the choice $t = 2$, $r = K/2$, and $s = 0$ is admissible because r is odd. Here r is proportional to K, so the exponents in Corollary D.2 are linear in K under the respective requirements of (i)–(iii).

For $K \equiv 2 \pmod 4$ with $K \geq 6$, part (iv) gives the stronger depth-two bound $2^{\Omega(K^2 \log(n'/K))}$. At $m' = \Theta(K \log n')$ with a sufficiently large constant factor, this matches the DNF upper bound in Proposition E.2 up to constants in the exponent whenever $\log(n'/K) = \Theta(\log n')$.

No single constant $c > 0$ allows an admissible choice $r \geq cK$ for every even K. To see the obstruction, let $L := \operatorname{lcm}(1, \ldots, M)$ and $K := 2L$. If $t \leq M$, then $t \mid L$ and hence $t \mid K$. The identity $K = tr + s$ forces $t \mid s$, so admissibility requires $s = 0$. But then $r = K/t = 2(L/t)$ is even, a contradiction. Every admissible decomposition therefore has $t > M$ and $r \leq K/t < K/M$. Since M can be arbitrarily large, the source parameter cannot remain proportional to K under this lift's arithmetic conditions. A different lift with such a guarantee remains an open question (Section 5.2).

The quadratic dependence on r in part (iv) nevertheless gives a depth-two exponent linear in K for every sufficiently large even K: it suffices to choose $r \geq \sqrt{K}/2$.

Corollary D.3 (Depth-two lower bounds for every even parameter)**.** *For every even $K \geq 36$, every $n' \geq 8K$, and every $m' \geq CK \log n'$, where C is the universal constant in Corollary D.2,*

$$\operatorname{size}_2(k\text{-}\mathsf{XOR}_{n',m',K}) \geq 2^{\Omega(K \log(n'/K))}. \tag{116}$$

Proof. Put $h := \lfloor \sqrt{K} \rfloor \geq 6$. At least one $t \in \{h - 1, h, h + 1\}$ does not divide K. Otherwise their least common multiple would divide K, but consecutive integers are coprime and the outer two have gcd at most two, so

$$\operatorname{lcm}(h - 1, h, h + 1) \geq \frac{(h - 1)h(h + 1)}{2} > (h + 1)^2 > K.$$

Choose such a t, take the largest odd $r \le K/t$, and put $s := K - tr$. Then $s > 0$ and $t \nmid s$, so the lift is admissible. Since $t \le \sqrt{K} + 1$, we have $r > K/t - 2 > \sqrt{K} - 3 \ge \sqrt{K}/2$. In particular $r \ge 3$ and $r^2 \ge K/4$, so Corollary D.2(iv) gives Equation (116). □

D.3 A PARITY projection at the principal row scale

Arrange the source bits in rows and consider the parity of each row. The XOR of these row parities equals the parity of all source bits. We construct an XOR instance in which a choice of target columns specifies any odd string of row parities, and the selected XOR vanishes exactly when that string agrees with the source row parities. An accepting selection therefore exists precisely when the source has odd total parity. The projection uses $O(k \log(n/k))$ rows, the same scale as the subgraph projections.

Guessing block parities is a standard way to encode PARITY with auxiliary variables; see Emdin et al. [EKMS22, §1.1]. Here the guesses are represented by target column labels. Bhattacharyya et al. [BGGS16, §6] use related indicator and consistency coordinates in a reduction to k-VectorSum with $O(k \log n)$ rows. We give the details needed to realize the parity guesses as a literal projection with exactly k selected columns and the stated row budget.

Write $\mathsf{PARITY}_\ell(x) := \bigoplus_{i=1}^{\ell} x_i$. Håstad's theorem [Hås89, Theorem 1, §4] implies that there is an absolute $\ell_0 > 1$ such that, for fixed $d \ge 2$ and $\ell > \ell_0^d$, every depth-d PARITY circuit has at least

$$\exp(\Omega_d(\ell^{1/(d-1)})) \tag{117}$$

gates. His depth and size count AND/OR layers and gates, with negations at the inputs [Hås89, §2.1], as in our model (Section 2.2). At depth two, the classical PARITY bound is $2^{\ell-1}$ clauses for a CNF, and equally many terms for a DNF by duality; see, for example, Emdin et al. [EKMS22, Lemma 2]. Thus $\mathrm{size}_2(\mathsf{PARITY}_\ell) \ge 2^{\ell-1}$ for every $\ell \ge 2$, without a further lower threshold on the source length.

Theorem D.4 (PARITY projection with $O(k \log(n/k))$ rows)**.** *Let $k \ge 3$ be odd and let $n \ge 4k$. Put*

$$b := \left\lfloor \log \frac{n}{2k} \right\rfloor, \qquad B := (k-1)b + 1, \qquad \ell := (k-1)B. \tag{118}$$

At row count $m_0 := (k-1)(b+1) + 2 \le k \log(n/k)$, there is a projection

$$\mathsf{PARITY}_\ell \le_{\mathrm{proj}} k\text{-}\mathsf{XOR}_{n,m_0,k}. \tag{119}$$

Consequently, for every fixed depth $d \ge 2$ and every $m \ge m_0$, if $d = 2$ or $\ell > \ell_0^d$,

$$\mathrm{size}_d(k\text{-}\mathsf{XOR}_{n,m,k}) \ge \exp\Big(\Omega_d\Big((k^2 \log(n/k))^{1/(d-1)}\Big)\Big). \tag{120}$$

Under the same condition on d and ℓ, the bound holds at every row count $m \ge k \log(n/k)$.

Proof. Arrange the $\ell = (k-1)B$ source variables as a matrix $(x_{h,t})_{h \in [B],\, t \in [k-1]}$, and let $p_h := \bigoplus_{t \in [k-1]} x_{h,t}$ be the parity of source row h. The vector $p \in \{0,1\}^B$ has odd parity exactly when the entire source input does.

The target has one auxiliary column $a_\star$ and $k-1$ blocks, each containing a column (t, v) for every label $v \in \{0,1\}^b$. These use $n_0 := 1 + (k-1)2^b < n$ positions, since $2^b \le n/(2k)$ and $n \ge 4k$. Fill the remaining positions with padding columns. Add one padding row equal to 1 on every padding column and 0 elsewhere, and put 0 in every other entry of each padding column. A zero-XOR selection must therefore use an even number of padding columns.

For every block t, add a selection row equal to 1 on that block and on $a_\star$, and 0 elsewhere. These are the selection and padding rows from the odd-XOR construction. Since k is odd, the same counting argument forces $a_\star$ and exactly one column from each block, with no padding selected. Thus the selections that pass these rows are exactly

$$\{a_\star\} \cup \{(t, u_t) : t \in [k-1]\}.$$

The labels $u_1, \dots, u_{k-1}$ specify $(k-1)b = B-1$ free bits. Append the unique final bit that makes their parity odd: for $u = (u_t)_{t\in[k-1]}$, define

$$y(u)_{(t-1)b+j} := u_{t,j} \quad (t \in [k-1],\ j \in [b]), \qquad y(u)_B := 1 \oplus \bigoplus_{t,j} u_{t,j}. \tag{121}$$

As the labels vary, the first $B-1$ coordinates vary freely and the last is determined by odd parity. Thus $y(u)$ ranges over exactly all odd strings in $\{0,1\}^B$.

It remains to impose $p = y(u)$ through data rows whose entries are constants or literals. Write $A_h(c)$ for the entry in data row h and column c. For $h < B$, write $h = (t_h - 1)b + j_h$ with $t_h \in [k-1]$ and $j_h \in [b]$. Define

$$\begin{aligned} A_h(a_\star) &:= 0, & A_h(t, v) &:= x_{h,t} \oplus \mathbf{1}[t = t_h] v_{j_h} && (h < B),\\ A_B(a_\star) &:= 1, & A_B(t, v) &:= x_{B,t} \oplus \bigoplus_{j\in[b]} v_j. \end{aligned}$$

The label v is fixed by the column index, so the bit added to $x_{h,t}$ is a constant. Each data entry is therefore a constant or a possibly negated source variable, and all selection and padding entries are constants. This verifies the literal requirement for a projection.

For a selection of the forced form, the XOR in data row h is $p_h \oplus y(u)_h$. Hence all data rows vanish precisely when

$$p_h = y(u)_h \qquad (h \in [B]). \tag{122}$$

If the source has odd total parity, then p is an odd string, so some tuple u realizes $y(u) = p$ and its selection passes every row. Conversely, any accepting selection has the forced form and satisfies Equation (122); since $y(u)$ is odd, so are p and the entire source input. This proves Equation (119) on all ℓ source variables.

There are $k-1$ selection rows, B data rows, and one padding row. Since $b + 1 \le \log(n/k)$ and $\log(n/k) \ge 2$, their total is

$$m_0 = k + B = (k-1)(b+1) + 2 \le (k-1)\log(n/k) + 2 \le k \log(n/k).$$

Appending zero rows reaches every $m \ge m_0$.

Finally, $n/k \ge 4$ ensures $b = \Theta(\log(n/k))$ despite the floor in Equation (118). Hence $\ell = (k-1)((k-1)b+1) = \Theta(k^2 \log(n/k))$. Projection monotonicity and (117) give Equation (120) when $\ell > \ell_0^d$; at $d = 2$, the classical depth-two bound above applies for every $\ell \ge 2$. □

Corollary D.5 (Tight depth-two bound for odd k-XOR)**.** *Let $k \ge 3$ be odd and $n \ge 4k$. For every $m \ge k \log(n/k)$,*

$$\mathrm{size}_2(k\text{-}\mathsf{XOR}_{n,m,k}) \ge 2^{\Omega(k^2 \log(n/k))}. \tag{123}$$

If moreover $m = O(k \log(n/k))$, then $\mathrm{size}_2(k\text{-}\mathsf{XOR}_{n,m,k}) = 2^{\Theta(k^2 \log(n/k))}$.

Proof. Apply the classical depth-two PARITY bound to the projection in Theorem D.4, with $\ell = \Theta(k^2 \log(n/k))$ and $\ell \geq 6$. Projection monotonicity and zero-row extension give Equation (123) in both depth-two orientations.

For the upper bound, Proposition E.2 gives a DNF with at most $1 + \binom{n}{k} 2^{(k-1)m}$ gates. If also $m = O(k \log(n/k))$, its logarithmic size is $O(km + k \log(en/k)) = O(k^2 \log(n/k))$. This matches the lower bound up to constants in the exponent. □

The number of PARITY variables is limited by the Satisfiability Coding Lemma of Paturi, Pudlák, and Zane [PPZ99, Lemma 4]: a CNF on ℓ variables with at most k literals per clause has at most $2^{\ell - \ell/k}$ isolated satisfying assignments. Here an assignment is isolated if flipping any one bit makes the formula false. Emdin et al. [EKMS22, §3.2, Lemma 6] use this bound to cover PARITY inputs by narrow CNFs. The following corollary applies the same covering argument to the sets of inputs accepted by individual supports, where a support is a set of k selected target indices.

Corollary D.6 (PARITY projection capacity)**.** *Let $n \geq k \geq 1$, and suppose that a projection $R : \{0,1\}^\ell \to (\mathbb{F}_2^m)^n$ satisfies*

$$k\text{-}\mathsf{XOR}_{n,m,k}(R(x)) = \mathsf{PARITY}_\ell(x) \qquad \text{for every } x \in \{0,1\}^\ell. \tag{124}$$

Then

$$\ell \leq k^2 \log(en/k) + k. \tag{125}$$

Proof. Fix a support S. Each of its row equations involves at most k source variables, so their conjunction can be written as a CNF with at most k literals per clause. By Equation (124), this CNF accepts only odd inputs, all of which are isolated. The cited bound therefore limits the accepted inputs to $2^{\ell - \ell/k}$. All $2^{\ell-1}$ odd inputs must be covered by the $\binom{n}{k}$ supports, giving

$$2^{\ell-1} \leq \binom{n}{k} 2^{\ell - \ell/k}. \tag{126}$$

Taking logarithms gives $\ell \leq k \log \binom{n}{k} + k \leq k^2 \log(en/k) + k$, as claimed. □

This recovers $\ell = O(k^2 \log(n/k))$ throughout the range of Theorem D.4, independently of m, so the construction uses the largest possible number of PARITY variables up to a constant factor. For fixed k in the range of Theorem 4.3 and $d \geq 3$, the PARITY bound is subpolynomial in n, while the subgraph bound is polynomial. In the range of Theorem 4.8, the subgraph Σ_3 exponent $\Omega(k \log(n/k))$ is also larger than the PARITY exponent $\Omega(k\sqrt{\log(n/k)})$. The subgraph framework additionally supplies the OV and SUM bounds.

For a concrete comparison, take odd $k = \Theta(\log^2 n)$ and $m = \Theta(k \log(n/k)) = \Theta(\log^3 n)$ with a sufficiently large constant factor. Then the construction uses $\ell = \Theta(\log^5 n)$ source variables, giving a lower bound $\exp(\Omega_d(\log^{5/(d-1)} n))$ at each fixed depth $d \geq 2$. Both the PARITY and subgraph projections fit within this row budget. Table 3 uses Theorem 4.6 for the subgraph bound at depth two, Theorem 4.8 for Σ_3, and Theorem 4.5 for the remaining rows. Under Conjecture 5.1, Corollary 5.2 gives the larger subgraph bound $\exp(\Omega_d(\log^3 n))$ also for Π_3 and fixed $d \geq 4$.

Class	PARITY route	Unconditional subgraph route	Larger bound
Depth 2	$\exp(\Omega_2(\log^5 n))$	$\exp(\Omega(\log^3 n))$	PARITY
Σ_3	$\exp(\Omega_3(\log^{5/2} n))$	$\exp(\Omega(\log^3 n))$	Subgraph
Π_3	$\exp(\Omega_3(\log^{5/2} n))$	$\exp(\Omega_3(\log^2 n))$	PARITY
Fixed $d \geq 4$	$\exp(\Omega_d(\log^{5/(d-1)} n))$	$\exp(\Omega_d(\log^2 n))$	Subgraph

Table 3: PARITY and subgraph lower bounds at odd $k = \Theta(\log^2 n)$ and $m = \Theta(k \log(n/k))$, with a sufficiently large constant factor. The depth-two PARITY bound is tight up to constants in the exponent (Corollary D.5).

E Upper Bounds and Comparison with Lower Bounds

We compare our lower bounds with circuits that enumerate witnesses. For SUM, guessing carries between short blocks of bits can reduce the checking cost. All sizes use the gate-count model of Section 2.2.

E.1 Enumerating witnesses

An OR over all possible witnesses allows us to focus on the cost of checking one fixed witness. OV and XOR have direct CNF checks, and XOR and SUM also have direct DNF checks. For SUM, guessing the carries gives another CNF check.

Checking a fixed witness. Even when the selected positions are fixed, checking XOR or SUM still contains the even-parity test. For XOR, let the first coordinate of each selected entry vary and set all other coordinates to zero. For SUM, let each selected entry be either 0 or 2^{m-1}; every pair of nonzero entries adds to zero modulo 2^m. In either case, the check accepts exactly when an even number of entries are nonzero. Since complementation preserves size and depth, the classical PARITY bound in Appendix D.3 requires $\exp(\Omega_d(k^{1/(d-1)}))$ gates to check either relation, for every $m \geq 1$, every fixed depth $d \geq 2$, and sufficiently large k depending on d.

Proposition E.1 (Elementary k-OV and k-XOR circuits)**.** *For all target parameters, the following top-disjunction depth-three bounds hold:*

$$\mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{OV}_{n,D,k}) \leq 1 + n^k(D+1), \tag{127}$$

$$\mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{XOR}_{n,m,k}) \leq 1 + \binom{n}{k}(1 + m2^{k-1}). \tag{128}$$

For OV this is $n^{k+o(1)}$ whenever $D = n^{o(1)}$, including fixed k and $D = O(k \log n)$; it is $n^{k+o(k)}$ if $k \to \infty$ and $\log D = o(k \log n)$. For XOR it is $(n/k)^{k+o(1)}$ if k is fixed, $n/k \to \infty$, and $\log m = o(\log(n/k))$; it is $(n/k)^{k+o(k)}$ if $k \to \infty$, $\log(n/k) \to \infty$, and $\log m = o(k \log(n/k))$.

Proof. For each OV tuple $(i_1, \ldots, i_k) \in [n]^k$, use the CNF $\bigwedge_{h \in [D]} \bigvee_{j \in [k]} \neg u_{j,i_j}(h)$: every coordinate must have a selected zero. For each XOR support $S \in \binom{[n]}{k}$, every row has an even-parity CNF with 2^{k-1} clauses, one excluding each odd assignment. Conjoin all m row CNFs. In each case, an OR over witnesses gives one top gate, one middle AND per witness, and the displayed number of bottom clause gates. The logarithmic estimates are

$$\log \mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{OV}) \leq k \log n + O(\log(D+1)), \qquad \log \mathrm{size}_{\Sigma_3}(k\text{-}\mathsf{XOR}) \leq k \log(n/k) + O(k) + \log m,$$

using $\binom{n}{k} \le (en/k)^k$. For OV, $D = n^{o(1)}$ makes the error $O(\log(D+1)) = o(\log n)$. If $k \to \infty$ and $\log D = o(k \log n)$, this error is $o(k \log n)$ instead.

For XOR with fixed k, the assumptions $n/k \to \infty$ and $\log m = o(\log(n/k))$ give $O(k) + \log m = o(\log(n/k))$. If $k \to \infty$, $\log(n/k) \to \infty$, and $\log m = o(k \log(n/k))$, then $O(k) + \log m = o(k \log(n/k))$. □

Proposition E.2 (Depth-two circuits for k-XOR)**.** *For all $n \ge k \ge 2$ and $m \ge 1$,*

$$\operatorname{size}_2(k\text{-}\mathsf{XOR}_{n,m,k}) \le 1 + \binom{n}{k} 2^{(k-1)m}. \tag{129}$$

Consequently,

$$\log \operatorname{size}_2(k\text{-}\mathsf{XOR}_{n,m,k}) \le (k-1)m + k \log(en/k) + O(1).$$

Proof. Fix a support $S = \{i_1 < \cdots < i_k\}$. Arbitrary values of the first $k-1$ selected m-bit words uniquely determine the last word making their XOR zero. There are $2^{(k-1)m}$ such assignments, and one AND gate on km literals recognizes each assignment. Taking the OR over these terms for all $\binom{n}{k}$ supports gives one top OR gate and $\binom{n}{k}2^{(k-1)m}$ term gates, proving Equation (129). This construction works for either parity of k. Taking logarithms and using $\binom{n}{k} \le (en/k)^k$ gives the logarithmic bound. □

At principal row counts, this DNF bound matches the odd-k depth-two lower bound in Corollary D.5.

Proposition E.3 (Elementary circuits for k-SUM)**.** *For all $n \ge k \ge 2$ and $m \ge 1$,*

$$\operatorname{size}_2(k\text{-}\mathsf{SUM}_{n,m,k}) \le 1 + \binom{n}{k} 2^{(k-1)m}. \tag{130}$$

Consequently,

$$\log \operatorname{size}_2(k\text{-}\mathsf{SUM}_{n,m,k}) \le (k-1)m + k \log(en/k) + O(1), \tag{131}$$

and the displayed depth-two circuit has size $2^{\Theta(km)}$ whenever $m \ge \log(en/k)$. There is also a top-disjunction depth-three block-carry circuit: for every integer $b \in [m]$, with $t := \lceil m/b \rceil$,

$$\operatorname{size}_{\Sigma_3}(k\text{-}\mathsf{SUM}_{n,m,k}) \le 1 + \binom{n}{k} k^{t-1}\left(1 + t2^{kb}\right). \tag{132}$$

In particular,

$$\log \operatorname{size}_{\Sigma_3}(k\text{-}\mathsf{SUM}_{n,m,k}) \le k \log(en/k) + O\left(\sqrt{km \log k} + k + \log m\right). \tag{133}$$

If $n \ge k^2$ and $m = \Theta(k \log(en/k))$, then

$$\operatorname{size}_{\Sigma_3}(k\text{-}\mathsf{SUM}_{n,m,k}) \le (n/k)^{O(k)}. \tag{134}$$

Proof. For a fixed support, arbitrary values of $k-1$ residues uniquely determine the last residue making the sum zero. Recognizing each of the $2^{(k-1)m}$ assignments by a conjunction of literals, then taking the OR over all supports, proves Equation (130). The binomial estimate gives Equation (131); when $m \ge \log(en/k)$ this particular DNF has $2^{\Theta(km)}$ gates, since even one support contributes $2^{(k-1)m}$ term gates.

The depth-three construction uses ordinary addition in blocks. Fix a support and split each selected word into $t = \lceil m/b \rceil$ consecutive blocks, starting at the least significant end. All blocks

Target	Transferred lower bound	Elementary Σ_3 upper bound
k-OV	$n^{\beta(k-1)}$	$n^{k+o(1)}$
k-XOR	$(n/k)^{\beta(k-1)}$	$(n/k)^{k+o(1)}$
k-SUM	$(n/k)^{\beta(k-1)}$	$(n/k)^{O(k)}$

Table 4: Fixed-k comparisons in the Σ_3 orientation. Dimensions and widths are within constant factors of the principal budgets (Lemma 2.4), with constants large enough for Theorem 4.3. Direct XOR requires odd k. Each pair agrees up to constants in the exponent.

have b bits except possibly the last. The initial carry is zero. Every intermediate carry stays in $\{0,\dots,k-1\}$: k block values of at most 2^b-1 plus an incoming carry of at most $k-1$ total at most $k2^b-1$, so the outgoing carry is again at most $k-1$. These carries are the only information passed between blocks. Knowing them would let us check all blocks independently, so we enumerate their k^{t-1} possible sequences. The final outgoing carry is discarded because the sum is taken modulo 2^m.

For each guessed sequence, check that every block's output bits are all zero and that every nonfinal block passes on the guessed carry. With the carries fixed, each check depends on at most kb input bits. Exclude each failing assignment by one clause to obtain a CNF with at most 2^{kb} clauses. Conjoin the clauses from all blocks under one AND gate, and take the OR over supports and carry sequences. A zero modular sum accepts on its true carry sequence. Conversely, in an accepting branch, induction from the least significant block, whose incoming carry is zero, shows that every guessed carry is the true one. Thus every output bit of the sum is zero. There are $\binom{n}{k}k^{t-1}$ branches, each with one AND gate and at most $t2^{kb}$ clause gates, proving Equation (132).

Taking logarithms gives

$$\log \mathrm{size}_{\Sigma_3}(k\text{-SUM}) \le k\log(en/k) + t\log k + kb + O(\log m). \tag{135}$$

Wider blocks mean fewer carries to guess but more input bits in each block check. Balance the costs $(m/b)\log k$ and kb by taking $b = \lceil\sqrt{m\log k/k}\rceil$, which lies in $[m]$ since $0 < \log k \le k$ and $m \ge 1$. Using $t \le m/b+1$ gives $t\log k + kb = O(\sqrt{km\log k}+k)$, proving Equation (133).

At principal width, $n \ge k^2$ gives $\log k \le \log(n/k)$. Substituting $m = \Theta(k\log(en/k))$ into Equation (133) bounds every term by $O(k\log(n/k))$, proving Equation (134). □

At the same budget scales, the upper bounds become $n^{k+o(k)}$, $(n/k)^{k+o(k)}$, and $(n/k)^{O(k)}$ in the growing-k polynomial range of Theorem 4.8. Thus the Σ_3 upper and lower exponents are linear in k for all three targets, with odd k for direct XOR.

E.2 Depth and width for k-SUM

The block-carry circuit matches the lower bound at the principal width, but its exponent can grow with larger widths. Under the host-size and width conditions for general SUM bounds, the lower exponent has scale $\min\{k\log(n/k), m\}$, while Equation (133) bounds the upper exponent by a constant multiple of

$$U_\Sigma(n,m,k) := k\log(en/k) + \sqrt{km\log k} + k + \log m. \tag{136}$$

At $n \ge k^2$ and principal width $m = \Theta(k\log(en/k))$, this scale is $\Theta(k\log(n/k))$. With a sufficiently large width constant, Theorem 4.8 gives a matching Σ_3 lower bound for sufficiently large k and $n \ge k^C$, with a suitable $C \ge 2$.

At smaller widths, $m \le k \log(n/k)$, enumerating supports may cost more than the lower exponent limited by m. At larger widths, the lower exponent stops growing at $k \log(n/k)$, and the ratio is

$$\frac{U_\Sigma(n,m,k)}{k \log(n/k)} = O\left(1 + \frac{\sqrt{m \log k/k}}{\log(n/k)} + \frac{k + \log m}{k \log(n/k)}\right). \tag{137}$$

The carry contribution to this ratio grows beyond a constant only when $m = \omega(k \log^2(n/k)/\log k)$. It is smaller than the direct depth-two DNF's $\Theta(m/\log(n/k))$ ratio, but can still grow with the width.

The block-carry circuit has orientation Σ_3; the depth-two and Π_3 gaps remain open.

F Proofs for the Depth-Three Bounds

We prove the single-CNF minterm bound of Theorem 4.7 and the circuit lower bounds of Theorem 4.8. The argument has three parts: a bound for one CNF, uniqueness of a planted copy in the random background, and a union bound over the middle CNFs of a circuit.

F.1 The single-CNF minterm bound

The revealing process is related to the decision-tree proofs of the switching lemma [Raz95, Bea94]; the use of a planted copy in a random background draws on the subgraph lower-bound framework of [Ros08, LRR17]. We combine these viewpoints to bound the probability that one fixed CNF needs every planted block for acceptance.

We use the random experiment of Section 4.4: for a simple loopless pattern F and host size $N \ge 1$, the background G includes each host-edge variable independently with probability $p \in (0,1]$, and the labels $\phi^*(v) \in [N]$ are independent and uniform, also independently of G. They determine the planted copy F^* in (64). As before, an input is represented by its set of variables assigned value 1.

Fix a nonempty independent set $I \subseteq V(F)$ of positive-degree vertices, and write $L := |I|$. We prove a stronger version of the single-CNF bound by grouping planted edges into blocks and requiring every block to be needed for acceptance. The full minterm event $\mathcal{M}(T)$ from Theorem 4.7 requires every individual planted edge to be needed, so it implies this block event.

We expose the labels outside I first. Since I is independent, every neighbor of a vertex $v \in I$ lies outside I and now has a fixed label. Thus choosing the one label at v determines all planted edges incident to v. For each candidate label $i \in [N]$ at v, the corresponding block is

$$B_{v,i} := \{X_{f,i,\phi^*(u)} : f = (v,u) \in E(F)\} \cup \{X_{f,\phi^*(u),i} : f = (u,v) \in E(F)\}. \tag{138}$$

The block $B_{v,i}$ contains all edges contributed by assigning label i to v, with its neighbors' planted labels held fixed. Revealing $\phi^*(v)$ selects one of these N candidate blocks. All candidate blocks are determined by the outside labels $\phi^*_{\overline{I}} := \phi^*|_{V(F)\setminus I}$. They are pairwise disjoint: two vertices of I control different pattern-edge groups, and two labels at the same vertex give different host-edge variables. This remains true when two vertices of I share a neighbor.

Let $F^*_{\overline{I}}$ be the planted edges not incident to I. These are already determined by the outside labels. Every other planted edge belongs to exactly one block $B_{v,\phi^*(v)}$. If $A \subseteq I$ records the vertices whose labels have been revealed, the background together with the partial planting is

$$H_A := G \cup F^*_{\overline{I}} \cup \bigcup_{v \in A} B_{v,\phi^*(v)}. \tag{139}$$

Thus $H_\varnothing = G \cup F_I^*$ adds only the outside planting to the background, and $H_I = G \cup F^*$ is the full planting. For a fixed CNF T, with arbitrary positive and negative literals, let $\mathcal{M}_I(T)$ be the *block minterm event*: T accepts only after every planted block has been added,

$$T(H_I) = 1 \qquad \text{and} \qquad T(H_A) = 0 \quad \text{for every } A \subsetneq I. \tag{140}$$

Lemma F.1 (Block minterm bound)**.** *Under the hypotheses of Theorem 4.7, with H_A as in (139), every fixed CNF T with $s := s(T) \geq 1$ clauses satisfies*

$$\Pr[\mathcal{M}_I(T)] \leq \left(\frac{C_0\big(L + \log(\mathrm{e}s)\big)}{pN} \right)^L, \tag{141}$$

with the same constant C_0 as in Theorem 4.7.

We prove Lemma F.1 using a process that reveals the planted labels, together with the auxiliary moment bound in Lemma F.2. Assume $s \geq 1$. The Lth power in (141) will come from the L successive steps needed to reveal every planted label in I.

Revealing the planted labels. Fix an ordering of the clauses of T. The CNF stays fixed throughout the process; adding a block changes the input by setting the block's variables to 1. For example, the clause $x \vee \neg y$ is false when $x = 0$ and $y = 1$; adding a block containing x sets x to 1 and makes the clause true. At a partial planting H_A, the process chooses the first false clause and tries to satisfy it by adding an unrevealed planted block. It starts with $A = \varnothing$ and makes uniform vertex choices independently of G and ϕ^*:

1. If $A = I$, declare success: all L labels have been revealed. Otherwise, if $T(H_A) = 1$, fail.
2. Otherwise $A \subsetneq I$ and $T(H_A) = 0$; call this state proper, and let C_A be the first clause false on H_A.
3. Choose $v \in I \setminus A$ uniformly and reveal $\phi^*(v)$. If C_A contains a positive literal whose variable lies in $B_{v,\phi^*(v)}$, add this block by replacing A with $A \cup \{v\}$ and repeat; otherwise fail.

Success means that all L labels have been revealed through permitted advances. The input H_A, and hence the selected clause C_A, depends only on the background, the outside labels, and the labels already revealed at A. No step inspects an unrevealed label before choosing its vertex.

We next count the choices that can turn a false clause true by setting one of its positive variables to 1. Fix the outside labels, so that every $B_{v,i}$ is determined before sampling G or exposing the labels in I. For a clause C and a state $A \subseteq I$, let $R_A(C)$ list the unrevealed vertex-label pairs whose blocks contain a positive literal of C:

$$\begin{aligned} R_A(C) &:= \big\{(v,i) \in (I \setminus A) \times [N] : \\ &\qquad\qquad C \text{ contains a positive literal whose variable lies in } B_{v,i}\big\}, \\ r_A(C) &:= |R_A(C)|. \end{aligned} \tag{142}$$

Thus $r_A(C)$ counts possible vertex-label choices for satisfying C when it is false. For each block containing a positive literal of C, choose one such literal $x_{C,v,i} \in C \cap B_{v,i}$ using only C and the fixed outside labels, before sampling G. The representatives use distinct variables because the blocks are disjoint. If any representative is present in G, its clause is already satisfied on every $H_A \supseteq G$. Consequently, a clause false on H_A must have all these representatives absent from G.

Call (A, C) *eligible* if every representative indexed by $R_A(C)$ is absent from G, and put

$$Z := \max\bigl\{r_A(C) : (A, C) \text{ eligible}\bigr\}. \tag{143}$$

Every pair consisting of a state A and a clause false on H_A is eligible, so Z bounds the number of vertex-label choices that can make the selected false clause true. Eligibility is only a necessary condition for falsity; it is useful because it can be checked without any label inside I. The maximum ranges over all states and clauses, so the bound still applies when the process chooses its clause using earlier observations. It is well defined since $R_I(C) = \varnothing$ for every clause C. With the outside labels fixed, Z depends only on G and is independent of the labels inside I.

Fix G and the outside labels, thereby fixing Z. We claim that

$$\frac{\Pr[\mathcal{M}_I(T) \mid G, \phi_I^*]}{L!} \;\le\; \Pr[\text{success} \mid G, \phi_I^*] \;\le\; \frac{Z^L}{L!\, N^L}. \tag{144}$$

These probabilities average over the labels in I; the success probability also averages over the vertex choices.

For the left inequality, step 3 is equivalent to sampling one of the $L!$ orders of I uniformly in advance and following it until the process stops. Fix all planted labels and suppose $\mathcal{M}_I(T)$ occurs. At every proper state, C_A is false on H_A but true on H_I. Every negative literal false on H_A has its variable already set to 1 and stays false as blocks are added. Thus some positive variable of C_A is 0 on H_A and becomes 1 in a remaining planted block. Choosing that block advances the process. Since every proper H_A rejects, repeating this choice constructs a successful order. Its probability is $1/L!$, so the success probability for each fixed planting is at least $\mathbf{1}[\mathcal{M}_I(T)]/L!$. Averaging over the labels in I proves the left inequality.

For the right inequality, condition on a history reaching a proper state. The history records only earlier vertex choices and their revealed labels, and all its tests used only those labels. The $q := |I \setminus A|$ unrevealed labels therefore remain mutually independent and uniform. The next vertex-label pair is consequently uniform among qN possibilities. Since C_A is false on $H_A \supseteq G$, all its representatives are absent from G, so (A, C_A) is eligible and $r_A(C_A) \le Z$. Step 3 therefore gives

$$\Pr[\text{advance from } A \mid G, \phi_I^*, \text{history}] = \frac{r_A(C_A)}{qN} \le \frac{Z}{qN}. \tag{145}$$

This bound holds after every possible history because Z was defined without any inside label. Success requires all L advances; multiplying the successive conditional upper bounds for $q = L, L-1, \dots, 1$ proves the right inequality in (144).

It remains to bound the Lth moment of Z. The next lemma considers all pairs (A, C) at once: there are 2^L subsets $A \subseteq I$ and s clauses C, hence at most $2^L s$ pairs.

Lemma F.2 (Tail and moment of Z)**.** *With the outside labels fixed, let* Pr *and* $\mathbb{E}$ *be taken over G alone. For every real $t \ge 0$,*

$$\Pr[Z \ge t] \le \min\{1, 2^L s\, \mathrm{e}^{-pt}\}. \tag{146}$$

Consequently, for an absolute constant $C_0 > 0$,

$$\mathbb{E}[Z^L] \le \left(\frac{C_0\bigl(L + \log(\mathrm{e}s)\bigr)}{p}\right)^L. \tag{147}$$

Both bounds are uniform in the fixed outside labels.

Proof. For a fixed pair (A, C) with $r_A(C) = r$, eligibility requires r distinct variables to be absent from G. Since the representatives were chosen before G, this has probability $(1-p)^r \le \mathrm{e}^{-pr}$. If $Z \ge t$, some eligible pair has $r \ge t$. A union bound over at most $2^L s$ pairs gives Equation (146).

Put $\alpha := \ln(2^L s) = L \ln 2 + \ln s$. In (146), the factor $2^L s$ equals e^{α}. Substituting $t = (\alpha + \lambda)/p$ therefore gives $2^L s\, \mathrm{e}^{-pt} = \mathrm{e}^{-\lambda}$, so

$$\Pr[pZ \ge \alpha + \lambda] \le \mathrm{e}^{-\lambda} \qquad (\lambda \ge 0). \tag{148}$$

Thus exceeding the threshold α/p by an additional λ/p has probability at most $\mathrm{e}^{-\lambda}$. This explains both the threshold $(L \ln 2 + \ln s)/p$ and the exponential decay beyond it.

To bound the Lth moment directly, use $w^L = \int_0^w L t^{L-1}\, dt$ for $w \ge 0$. Taking expectations expresses the moment as an integral of upper-tail probabilities. Bound the part below α by α^L and use (148) above α to obtain

$$\begin{aligned}
\mathbb{E}[(pZ)^L] &= \int_0^\infty L t^{L-1} \Pr[pZ > t]\, dt \\
&\le \alpha^L + \int_0^\infty L(\alpha + u)^{L-1} \mathrm{e}^{-u}\, du \\
&= \int_0^\infty (\alpha + u)^L \mathrm{e}^{-u}\, du.
\end{aligned}$$

The last equality is integration by parts. To bound the remaining integral, expand $(\alpha + u)^L$ by the binomial theorem. Repeated integration by parts gives $\int_0^\infty u^j \mathrm{e}^{-u}\, du = j!$, and $j! \le L^j$ for $0 \le j \le L$ because each of its j factors is at most L. Hence

$$\mathbb{E}[(pZ)^L] \le \sum_{j=0}^{L} \binom{L}{j} \alpha^{L-j} j! \le \sum_{j=0}^{L} \binom{L}{j} \alpha^{L-j} L^j = (\alpha + L)^L.$$

Finally, $\alpha + L = (1 + \ln 2)L + \ln s \le 2\big(L + \log(\mathrm{e}s)\big)$, since $1 + \ln 2 < 2$ and our base-two logarithm satisfies $\ln s \le \log s$ for $s \ge 1$. Dividing the moment bound by p^L proves Equation (147), with a constant independent of the outside labels. □

Proof of Lemma F.1 and Theorem 4.7. Multiplying the outer inequality in (144) by $L!$ bounds the conditional minterm probability by Z^L/N^L. Averaging over G and applying Lemma F.2 gives (141) for each fixed choice of outside labels. Its uniformity then permits averaging over the outside labels as well.

Finally, a proper partial planting omits at least one whole planted block. Such a block is nonempty because its vertex has positive degree, and is disjoint from the other blocks and F^*_I. Thus the planted part of H_A is a proper subset of F^* whenever $A \subsetneq I$. The full minterm event therefore implies the block minterm event, $\mathcal{M}(T) \subseteq \mathcal{M}_I(T)$. Applying (141) gives the full minterm bound (65), proving Theorem 4.7. □

F.2 Uniqueness of the planted copy

We now show that a sufficiently sparse background makes the source function accept the full planting and reject every proper deletion with high probability. It suffices to show that the planted copy is unique and none of its edges already belongs to the background. Another colored copy would choose different host labels for some pattern vertices; regularity forces these choices to require many edges from the sparse background. The probabilistic argument uses the regular core, and the matching padding is restored by projection when proving Theorem 4.8.

Lemma F.3 (A unique planted copy with no background overlap)**.** *Fix an integer $\Delta \geq 3$ and a constant $\theta > 2/\Delta$. Let $F = F_k$ be a simple loopless Δ-regular graph with $v(F) = O(k)$, and choose the background density*

$$p := N^{-\theta}. \tag{149}$$

There is a constant $C_{\Delta,\theta} > 0$ such that, uniformly for $N \geq k^{C_{\Delta,\theta}}$, the following event has probability $1 - o(1)$ as $k \to \infty$: no variable of the planted copy F^* already belongs to G, and F^* is the unique F-copy in $G \cup F^*$, where an F-copy means an accepting map $\psi : V(F) \to [N]$ in the sense of Definition 2.1. We call this event the good event*. On this event, $G \cup R$ is a 0-input of F-SUB$_N$ for every proper subset $R \subsetneq F^*$.*

Proof of Lemma F.3. Fix the planted labels ϕ^*. All probability estimates below are over G and uniform in this choice. The planted copy is present in $G \cup F^*$ by construction. Since it has $e(F) = \Delta v(F)/2 = O(k)$ edges, a union bound gives

$$\Pr[F^* \cap G \neq \varnothing] \leq e(F)p = o(1) \tag{150}$$

uniformly for $N \geq k^{C_{\Delta,\theta}}$ once $C_{\Delta,\theta}$ is sufficiently large: the bound is $O(k^{1-C_{\Delta,\theta}\theta})$.

It remains to rule out another colored copy. Such a copy chooses a host label $\psi(v)$ for each pattern vertex v and requires every corresponding edge variable to be present, as in Definition 2.1. We compare this assignment with ϕ^* by setting $J := \{v : \psi(v) \neq \phi^*(v)\}$ and $j := |J| \geq 1$. Outside J, the labels agree with the planting; for each fixed J, there are at most N^j choices for the labels inside it.

Every pattern edge meeting J has a changed endpoint label, so the variable it requires differs from the planted variable for that edge. It also differs from all other planted variables, since each variable is indexed by its pattern edge. Thus every such edge must be supplied by G. Distinct pattern edges require distinct variables, whose presence in G is independent.

Regularity now controls how many background edges are needed. The vertices of J have Δj edge-endpoints altogether. An edge internal to J accounts for two of them, while an edge leaving J accounts for one. Writing $e_F(J)$ and $e_F(J, V(F) \setminus J)$ for these two edge counts gives

$$\Delta j = 2e_F(J) + e_F(J, V(F) \setminus J),$$

so at least $\Delta j/2$ edges meet J. For a fixed assignment ψ, the probability that all its required background edges are present is therefore at most $p^{\Delta j/2}$. There are $\binom{v(F)}{j}$ choices of the changed vertex set and at most N^j assignments on it. Summing these probabilities bounds the expected number of other copies:

$$\begin{aligned} \mathbb{E}[\#\{\psi \neq \phi^* : \psi \text{ is a copy in } G \cup F^*\}] &\leq \sum_{j=1}^{v(F)} \binom{v(F)}{j} N^j p^{\Delta j/2} \\ &\leq \sum_{j \geq 1} \big(v(F) N^{-(\theta\Delta/2-1)}\big)^j = o(1). \end{aligned} \tag{151}$$

The second inequality uses $\binom{v(F)}{j} \leq v(F)^j$ and $p = N^{-\theta}$. Since $\theta > 2/\Delta$, the power $\theta\Delta/2 - 1$ is positive. Choose $C_{\Delta,\theta}$ so that $C_{\Delta,\theta}(\theta\Delta/2 - 1) > 2$ and $C_{\Delta,\theta}\theta > 2$. At the smallest permitted host size $N = k^{C_{\Delta,\theta}}$, the geometric ratio is $O(k^{1-C_{\Delta,\theta}(\theta\Delta/2-1)}) = O(k^{-1})$, and the overlap probability bound is also $O(k^{-1})$. Both decrease as N increases, so both bounds are $o(1)$ uniformly for $N \geq k^{C_{\Delta,\theta}}$.

Markov's inequality bounds the probability of another copy by its expected count. Combining this with Equation (150) proves that the good event has probability $1 - o(1)$, still uniformly after averaging over the planted labels.

Finally, let $R \subsetneq F^*$ on the good event. Any copy in $G \cup R$ would also lie in $G \cup F^*$, so uniqueness would force it to be the planted copy. But every edge in $F^* \setminus R$ is absent from R and, by disjointness, from G. Thus the background cannot replace a deleted planted edge, and $G \cup R$ contains no copy. It is therefore a 0-input of F-SUB$_N$. □

F.3 Proof of the depth-three circuit bounds

We combine the single-CNF probability bound with the good event from Lemma F.3. We first explain the choice of degree and density, then prove the source bound and transfer it to the three targets.

The density must serve two purposes: a larger p strengthens the single-CNF bound through its denominator pN, while a smaller p prevents other copies. For a bipartite Δ-regular core with $k - O(1)$ edges, one side has $L = k/\Delta + O(1)$ vertices. The main factor in (65) is $N^{-(1-\theta)L}$; a sufficiently large polynomial host absorbs the numerator with a small loss in the exponent. Taking θ just above $2/\Delta$ therefore gives limiting source coefficient $(1/\Delta)(1 - 2/\Delta)$. Increasing Δ permits a denser background but leaves fewer vertices on the independent side. The coefficient is largest at $\Delta = 4$, where it equals $1/8$.

We now use the bipartite 4-regular core $F = F_k^{\exp}$ of Lemma 2.2 and one side $I = I_k$, with $L := |I_k|$. By Equation (4), we have $4L = e(F_k^{\exp}) = k - O(1)$ and hence $L = k/4 + O(1)$. Since $v(F) = e(F)/2 \le (k-1)/2$, Lemma F.3 applies with $\Delta = 4$ whenever $\theta > 1/2$, using the density (149). We write $C_{4,\theta}$ for its good-event threshold constant. Choosing θ sufficiently close to $1/2$ gives every source coefficient below $1/8$ by the tradeoff above. The proof uses this independent side and regularity; expansion and κ play no role.

Proof of Theorem 4.8. Fix $0 < \gamma < 1/8$. Our uniqueness bound requires $\theta > 1/2$, while the union bound must leave the main single-CNF exponent $(1-\theta)k/4$ larger than γk. We therefore choose

$$\frac{1}{2} < \theta < 1 - 4\gamma. \tag{152}$$

Thus $\gamma < (1-\theta)/4$. We reserve room for two losses in the single-CNF bound (65): replacing the number of blocks L by a lower bound $(1/4 - \eta)k$, and bounding its numerator $C_0(L + \log(es))$ by N^η for the clause counts used below. The resulting exponent per block will be $1 - \theta - \eta$, so we need $\gamma < (1/4 - \eta)(1 - \theta - \eta)$. Since this inequality is strict at $\eta = 0$, choose $\eta > 0$ small enough that

$$\eta < \min\left\{\frac{1}{4}, 1 - \theta\right\} \qquad \text{and} \qquad \gamma < \left(\frac{1}{4} - \eta\right)(1 - \theta - \eta). \tag{153}$$

By Equation (4), after increasing the threshold in k we have $L \ge (1/4 - \eta)k$.

Suppose that a Σ_3 circuit $\mathcal{C} = \bigvee_{i=1}^{t} T_i$ of size $S \le N^{\gamma k}$ computes $F_k^{\exp}$-SUB$_N$. Put $M_{\mathrm{src}} := e(F_k^{\exp})N^2$ and $\widehat{S} := S + 2M_{\mathrm{src}}$. The normal-form bound (6) gives $t \le \widehat{S}$ and $s(T_i) \le \widehat{S}$ for every i, including the allowance for literals that bypass gates. All the CNFs T_i are fixed before the random input is sampled.

On the good event, correctness gives $\mathcal{C}(G \cup F^*) = 1$ and $\mathcal{C}(G \cup R) = 0$ for every proper $R \subsetneq F^*$. Some T_i accepts the full planting. Since an OR is 0 only when all its inputs are 0, every T_i rejects every proper deletion. In particular, the same CNF that accepts the full planting satisfies $\mathcal{M}(T_i)$.

Thus the good event is contained in the union of the full minterm events of the middle CNFs. By Theorem 4.7 and a union bound,

$$\Pr[\text{the good event}] \le \widehat{S}\left(\frac{C_0\big(L+\log(\mathrm{e}\widehat{S})\big)}{pN}\right)^L. \tag{154}$$

It remains to make this upper bound tend to zero. Choose $C_E(\gamma)$ large enough to imply the good-event threshold $N \ge k^{C_{4,\theta}}$ and to satisfy $C_E(\gamma) > 2/\eta$. Then $N \ge k^{C_E(\gamma)}$ gives $k \le N^{\eta/2}$.

The allowance for input literals is quadratic in N: $2M_{\mathrm{src}} \le 2(k-1)N^2$. After increasing $k_E(\gamma) \ge k_0$, it is at most $N^{\gamma k}$ throughout the polynomial-host range, so $\widehat{S} \le 2N^{\gamma k}$.

Consequently $L + \log(\mathrm{e}\widehat{S}) = O(k \log N)$. There is therefore a constant A_γ such that $C_0(L + \log(\mathrm{e}\widehat{S})) \le A_\gamma k \log N$. Since $\log N = o(N^{\eta/2})$, increasing the threshold in k ensures $A_\gamma \log N \le N^{\eta/2}$ for every allowed host size. Combining this with the host bound above gives

$$C_0(L + \log(\mathrm{e}\widehat{S})) \le N^{\eta/2} N^{\eta/2} = N^{\eta}.$$

We can now bound each single-CNF probability by a power of N:

$$\left(\frac{C_0(L+\log(\mathrm{e}\widehat{S}))}{pN}\right)^L \le \left(\frac{N^\eta}{N^{1-\theta}}\right)^L = N^{-(1-\theta-\eta)L}.$$

The union bound (154) multiplies this by at most $\widehat{S} \le 2N^{\gamma k}$, giving

$$2N^{\gamma k - (1-\theta-\eta)L} \le 2N^{[\gamma-(1-\theta-\eta)(1/4-\eta)]k} = o(1). \tag{155}$$

The exponent is negative by Equation (153). Hence the assumed circuit would make the good event have probability $o(1)$, contradicting Lemma F.3, which gives probability $1 - o(1)$. This proves $\mathrm{size}_{\Sigma_3}(F_k^{\exp}\text{-SUB}_N) > N^{\gamma k}$.

We now restore the matching padding needed for exactly $k-1$ edges. Setting every padding-edge variable to 1 and leaving core variables unchanged gives $F_k^{\exp}\text{-SUB}_N \le_{\mathrm{proj}} P_k^{\exp}\text{-SUB}_N$ (Equation (95)). An accepting core copy extends over these always-present matching edges, so projection monotonicity within Σ_3 gives the first inequality in Equation (66).

The square-root host substitutions explain the factor of two between source and target coefficients. Fix $0 < c < \gamma/2$ and choose $C'(\gamma, c) > \max\{2C_E(\gamma) + 3, 3\}$. For $n \ge k^{C'(\gamma,c)}$, the host estimates of Lemma 2.4 give both N_{OV} and $N_{\oplus,\Sigma}$ at least $k^{C_E(\gamma)}$ once k is sufficiently large. This range also gives $n \ge 16k$ for the lemma and $n \ge k^3$ for the sharper SUM budget (18).

Throughout the allowed range $n \ge k^{C'(\gamma,c)}$, we have $n/k \ge k^{C'(\gamma,c)-1}$. Since $C'(\gamma, c) > 1$ and $\gamma/2 - c > 0$,

$$(n/k)^{\gamma/2-c} \ge k^{(C'(\gamma,c)-1)(\gamma/2-c)} \longrightarrow \infty.$$

Thus we may choose the lower threshold $k_1(\gamma, c)$ large enough that $k \ge k_E(\gamma)$ and $(n/k)^{\gamma/2-c} \ge 2^\gamma$ for every allowed pair (k, n). This verifies (20) for the smaller base n/k, and therefore also for OV's base n. Apply Lemma 2.4 with source exponent $a = \gamma k$ to obtain all three target bounds within Σ_3, with oddness needed only for XOR. Every fixed $0 < c < 1/16$ is attained by choosing $2c < \gamma < 1/8$. □

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