Electronic Colloquium on Computational Complexity, Report No. 330 (2026)

Improved Bounds for Depth-Two Threshold Circuits Computing Andreev's Function

Christian Reitwießner*

November 2026

Abstract

Andreev's function reads a bit from a truth table at an address given by the parities of blocks of input bits. Kane and Williams used it for the first superlinear gate lower bounds against depth-two threshold circuits and asked for tight bounds. We improve the bounds by showing:

- an exact circuit with $(5/16 + o(1))n^2$ gates and polynomial integer weights (previously $O(n^3/\log n)$);
- for agreement on a $\frac{1}{2} + \varepsilon$ fraction of inputs, matching bounds of $\varepsilon^2 n^{3/2}$ gates up to polylogarithmic factors (for some ε, see below), in particular: an explicit circuit with $O(\varepsilon^2 n^{3/2}\sqrt{\log(1/\varepsilon)})$ gates, and a lower bound of $\Omega(\varepsilon^2 n^{3/2}/\log^2 n)$ (previously ε^3 and $\log^3 n$);
- worst-case lower bounds of $\Omega(n^{3/2}/\log^2 n)$ gates and $\Omega(n^{5/2}/\log^{5/2} n)$ wires (previously $\log^3 n$ and $\log^{7/2} n$ in the denominators).

In our circuits, and in that of Kane and Williams, each bottom gate's weights have a single sign within every address block. We show that such circuits for Andreev's function need $\Omega(n^2/\log^3 n)$ gates. Within this class our exact circuit is therefore optimal up to polylogarithmic factors, and beating n^2 by more than that requires mixed signs.

All numbered results are formalized in Lean; the asymptotic consequences are derived from explicit bounds in the parameters.

1 Setting and results

Lower bounds for depth-two threshold circuits are a classical frontier of circuit complexity: for explicit functions in P, nothing beyond roughly $n^{3/2}$ gates is known. The first superlinear gate lower bound is due to Kane and Williams [KW16], whose hard function is Andreev's function: look up a bit in a table, at an address given by parities. They proved $\Omega(n^{3/2}/\log^3 n)$ gates, gave an $O(n^3/\log n)$ circuit, and asked for asymptotically tight bounds. This article narrows the gap from both sides, settles the approximate version up to polylogarithmic factors, and locates what remains of the exact gap: if every bottom gate has weights of a single sign inside each address block, then $\Omega(n^2/\log^3 n)$ gates are necessary, so a substantially subquadratic circuit must have bottom gates with weights of both signs inside some address block.

*christian@reitwiessner.de

ISSN 1433-8092

Andreev's function. Fix integers $k, m \geq 1$. The input consists of

- a *table* x of 2^k bits x_c, one for each address $c \in \{0,1\}^k$, and
- k *address blocks* of m bits each; block i consists of the bits $a_{i,1}, \ldots, a_{i,m}$.

The function is

$$A(x,a) = x_b, \qquad \text{where } b_i = \left(\sum_{j=1}^{m} a_{i,j}\right) \bmod 2 \quad \text{for } i = 1, \ldots, k.$$

Write $N = 2^k$ for the length of the table, so the input length is $n = N + km$. In the *balanced case* the table and the address part have equal length, $km = N$, so $n = 2N$ and $k = \log_2(n/2)$. Exact balance requires k to be a power of two. Asymptotic statements refer to this balanced sequence; statements for general k, m do not require this assumption. We use log for the base-two logarithm and ln for the natural logarithm.

Threshold circuits. For a statement P, the bracket $[\,P\,]$ is 1 if P holds and 0 otherwise. A *threshold gate* with inputs $z_j \in \{0,1\}$, real weights ω_j and threshold τ computes $\left[\sum_j \omega_j z_j \geq \tau\right]$. In a *depth-two threshold circuit* ($\mathsf{THR} \circ \mathsf{THR}$) the *bottom* gates read only input bits, and a single *top* gate reads the bottom gates' output and input bits. Its *size* is its total number of gates, top gate included; its *wire count* is the number of connections with nonzero weight. For $0 < \varepsilon \leq 1/2$, a circuit *ε-agrees* with a Boolean function if it outputs the function's value on at least a $\frac{1}{2} + \varepsilon$ fraction of all inputs. Agreement fractions are measured under the uniform distribution on inputs.

Results and the state of the art. Stronger lower bounds against $\mathsf{THR} \circ \mathsf{THR}$ are known for functions in the much larger class E^{NP}; they are proved by the algorithmic method [Tam16, ACW16, CTW26]. For functions in P, which includes Andreev's function, the best known gate lower bounds are of order $n^{3/2}/\operatorname{polylog} n$ [KW16, KKL17], as also noted by Chen, Tal and Wang [CTW26]. Our bounds improve the explicit polylogarithmic factors and the dependence on ε.

Function	Resource	Bound	Best previously known	This article	
Functions in P					
Andreev, exact	gates	upper	$O(n^3/\log n)$ [KW16]	$(5/16 + o(1))n^2$	Thm. 5
		lower	$\Omega(n^{3/2}/\log^3 n)$ [KW16] $n^{3/2}/(\log n)^{O(1)}$ [KKL17]†	$\Omega(n^{3/2}/\log^2 n)$	Thm. 8
	wires	lower	$n^{5/2}/(\log n)^{O(1)}$ [KKL17]†	$\Omega(n^{5/2}/\log^{5/2} n)$	Thm. 8
Andreev, ε-agree	gates	upper	none stated	$O(\varepsilon^2 n^{3/2}\sqrt{\log(1/\varepsilon)})$	Thm. 7
		lower	$\Omega(\varepsilon^3 n^{3/2}/\log^3 n)$ [KW16]	$\Omega(\varepsilon^2 n^{3/2}/\log^2 n)$	Thm. 8
	wires	lower	$\Omega(\varepsilon^3 n^{5/2}/\log^{7/2} n)$ [KW16]	$\Omega(\varepsilon^2 n^{5/2}/\log^{5/2} n)$	Thm. 8
Andreev, sign-coherent, exact‡	gates	lower	none stated	$\Omega(n^2/\log^3 n)$	Thm. 13
Functions in E^{NP}					
some F, exact	gates	lower	$n^{2-\delta}$ [Tam16, ACW16] $n^{5/2-\delta}$ [CTW26]	–	
some F, ε-agree	wires	lower	$n^3/\log^{12} n$ [Nag26]	–	

Notes on the table:

- All rows concern $\mathsf{THR} \circ \mathsf{THR}$ circuits. The Andreev rows are asymptotic in the balanced case; its exact upper bound has the explicit finite version (2).
- The ε-agreement rows for Andreev's function hold for $\varepsilon \geq 3\sqrt{2}(\log_2 n + 1)/\sqrt{n}$. Below this, a single majority gate over the table already agrees on a $\frac{1}{2} + \Omega(1/\sqrt{n})$ fraction. The upper bound is stated for $\varepsilon \leq \frac{1}{8}$. For $\frac{1}{8} < \varepsilon < \frac{1}{2}$, use Part 1 of Theorem 7 with $\delta = \frac{1}{2} - \varepsilon$; at $\varepsilon = \frac{1}{2}$, use the exact construction.
- †For a variant of Andreev's function whose table is four times as long as the address part.
- ‡For circuits whose bottom gates are *sign-coherent*: inside every address block, the weights of a gate are all ≥ 0 or all ≤ 0 (Section 6). The circuit of Theorem 5 is of this kind.

The remaining gap for unrestricted Andreev circuits concerns exact computation: between $n^{3/2}/\log^2 n$ and $(5/16 + o(1))n^2$ gates. For circuits with sign-coherent bottom gates the gap is only polylogarithmic. For any fixed $\delta > 0$, a circuit with $n^{2-\delta}$ gates therefore needs, for sufficiently large n, bottom gates whose weights have mixed signs inside a block. The two polynomial-size constructions we know of avoid such gates: besides ours, the circuit of Kane and Williams [KW16, Thm. 1.2] builds each of its parities from gates that threshold the plain sum of the parity's variables [HMPST93, Prop. 2.1], so it is sign-coherent as well.

Related work. Andreev's function comes from formula complexity: for De Morgan formulas, shrinkage gives it a lower bound of $n^{3-o(1)}$, the best known for an explicit function [And87, Hås98, Tal14]. For depth-two threshold circuits, exponential lower bounds for inner product hold when the top gate has small weights [HMPST93, Nis93]. With arbitrary weights at both levels, the bounds of Kane and Williams [KW16] were the first superlinear gate and superquadratic wire bounds; Kabanets, Kane and Lu [KKL17] and Chen, Santhanam and Srinivasan [CSS18] refine their restriction method and use it for satisfiability algorithms, and the algorithmic method gives stronger bounds for functions in E^{NP} [Tam16, ACW16, CTW26]. Sign-coherent gates are unate in each address block; unateness also drives the parity lower bounds of Impagliazzo, Paturi and Saks [IPS97].

How to read this article. Section 2 lists the new techniques and where they are used. Sections 3–4 (upper bounds), Section 5 (lower bounds) and Section 6 (sign-coherent circuits) can be read independently. Section 7 lists open questions.

2 New techniques

The proofs combine standard tools, namely Fourier expansion of the lookup, threshold staircases for functions of a linear form (as for symmetric functions [HMPST93, Prop. 2.1]), and the random restrictions of Kane and Williams [KW16], with the following new ingredients.

- **Removing the parity factor.** Used for the exact upper bound (Lemma 2, Proposition 4, Theorem 5). The Fourier expansion writes the lookup as a sum of products "table coefficient $\times$ address parity". For $\varphi(u) := (-1)^u u$ and even X, $X(-1)^t = \varphi(X+t) - \varphi(t)$, so every product becomes the difference of two functions of affine forms, each implemented by a threshold staircase. Kane and Williams build each of about $n^2/\log n$ parities separately. Collecting the

terms that share the same address parity $(-1)^{S\cdot b}$ leaves only N terms, each with an integer coefficient X_S, and saves a factor $n/\log n$. Sharing quadratic corrections across characters and absorbing errors at a few special addresses gives the constant 5/16.

- **Odd error functions are orthogonal.** Used for the upper bounds for approximate computation (Lemma 6, Theorem 7). For each Fourier coefficient, we pair two reflected forms and truncate them to the same symmetric window. Their combined error is an odd function of that coefficient. A symmetry of the uniform table flips one coefficient and fixes another, so the errors of different characters are orthogonal although the coefficients are not independent. Their squares add, and no union bound over characters is needed. Keeping only the characters of the first r address bits gives a majority vote over a subcube of the table and the ε^2 tradeoff.

- **Fixed-weight counting.** Used for both lower bounds, the general one and the one for sign-coherent circuits (Lemma 11, Theorems 8 and 13). After a restriction, a surviving gate keeps its original weights on the live bits; only its threshold moves. This fact is standard [KW16, CSS18, CTW26], but we use it for counting: a survivor costs $O(k + \log s)$ bits instead of the $\Theta(k^2)$ bits of an arbitrary threshold function [KW16, KKL17, CSS18]. This saves one logarithm. It works because the circuit is fixed before the random table is drawn.

- **A union bound uniform in the number of survivors.** Used for the lower bound for approximate computation (Lemma 12, Theorem 8). Instead of bounding the survivors by Markov's inequality, we bound the correlation for all survivor counts at once and average with Cauchy–Schwarz. This improves ε^3 to ε^2.

- **The cube restriction.** Used for the lower bound for sign-coherent circuits (Lemmas 14 and 15, Theorem 13). Instead of fixing the other address bits at random, fix a prefix of every block and let its length run through all values. On every resulting unit cube of block counts the circuit still reveals the whole table, while a sign-coherent gate moves monotonically through the grid and changes value at most once along each diagonal line, a classical fact for unate functions [IPS97, Lemma 2.1]. So it is non-constant on at most a k/m fraction of the cubes, compared with the $O(\sqrt{n}/m)$ survival bound for random restrictions.

3 Exact quadratic circuits

Notation for Sections 3 and 4. Let $|a_i|$ be the number of ones in block i, so the address is b with $b_i = |a_i| \bmod 2$. For $S \subseteq \{1,\dots,k\}$ and $c \in \{0,1\}^k$ write $S \cdot c = \sum_{i\in S} c_i$, and

$$X_S(x) \coloneqq \sum_{c\in\{0,1\}^k} (-1)^{S\cdot c + x_c}, \qquad t_S(a) \coloneqq \sum_{i\in S} |a_i|.$$

$X_S(x)$ is N times the Fourier coefficient at S of $c \mapsto (-1)^{x_c}$; it is even and lies in $[-N, N]$. $t_S(a)$ is the number of ones in the blocks in S, and lies in $[0, |S|m]$. We drop the arguments when they are clear. We write $[r] = \{1,\dots,r\}$, and $c_{[r]}$ for the first r coordinates of c.

Plan. Testing the sign of the score $N(-1)^{x_b}$ computes Andreev's function. Lemma 1 writes it as a sum of products, Lemma 2 removes the products, and Lemma 3 turns the result into threshold gates. We then refine this to a smaller score with the same sign, proving the 5/16 bound.

Lemma 1 (Lookup as a sum of products)**.** *For every table x and address input a,*

$$\sum_{S\subseteq\{1,\dots,k\}} X_S(x)\,(-1)^{t_S(a)} = N\,(-1)^{x_b}.$$

Proof. Since $|a_i| \equiv b_i \pmod 2$, we have $(-1)^{t_S} = (-1)^{S\cdot b}$. Exchanging the two sums, the left side is

$$\sum_c (-1)^{x_c} \sum_{S\subseteq\{1,\dots,k\}} (-1)^{S\cdot c + S\cdot b} = \sum_c (-1)^{x_c} \sum_{S\subseteq\{1,\dots,k\}} \prod_{i\in S} (-1)^{c_i+b_i}.$$

Choosing S means choosing, for each i, whether $i \in S$, so the inner sum factors as $\prod_{i=1}^{k}(1 + (-1)^{c_i+b_i})$. This product is N for $c = b$ and 0 otherwise. □

Lemma 2 (Removing the parity factor)**.** *Define φ on the integers by*

$$\varphi(\theta) \coloneqq (-1)^{\theta}\,\theta.$$

If X is an even integer and t is an integer, then $X(-1)^t = \varphi(X+t) - \varphi(t)$.

Proof. Because X is even, $X + t$ has the parity of t, so $\varphi(X+t) = (-1)^t(X+t)$. □

Lemma 3 (Threshold staircase)**.** *Let L be an integer-valued quantity with $\ell \le L \le h$, where ℓ, h are integers, and let $g : \mathbb{Z} \to \mathbb{Z}$. Then*

$$g(L) = g(\ell) + \sum_{\theta=\ell+1}^{h} \big(g(\theta) - g(\theta-1)\big)\big[\,L \ge \theta\,\big].$$

Proof. The brackets equal 1 exactly for $\theta \le L$, and the sum telescopes. □

When L is an affine form in the input bits, each $\big[\,L \ge \theta\,\big]$ is one threshold gate, so $g(L)$ costs one gate per value of θ, i.e. $h - \ell$ gates. Both t_S and $X_S + t_S$ are such forms, since $(-1)^{x_c} = 1 - 2x_c$:

$$X_S + t_S = \sum_c (-1)^{S\cdot c} - 2\sum_c (-1)^{S\cdot c} x_c + \sum_{i\in S}\sum_{j=1}^{m} a_{i,j}. \tag{1}$$

We now combine Lemmas 1–3 into a quadratic circuit for Andreev's function. Theorem 5 below refines it to $(5/16 + o(1))n^2$ gates, with the proof in Appendix A.

Proposition 4 (A first quadratic construction)**.** *For all $k, m \ge 1$, Andreev's function has an exact depth-two threshold circuit with $2N^2 + kmN$ bottom gates and one top gate. In the balanced case this is $\frac{3}{4}n^2 + 1$ gates. Bottom weights lie in $\{0, \pm 2\}$ on table bits and in $\{0, 1\}$ on address bits, and top weights are integers of absolute value at most $2n - 1$.*

Proof. By Lemmas 1 and 2, $N(-1)^{x_b} = \sum_S \big(\varphi(X_S + t_S) - \varphi(t_S)\big)$. Apply Lemma 3 to every term, with $X_S + t_S$ ranging over $[-N,\, N + |S|m]$ and t_S over $[0, |S|m]$. Put $d_\theta \coloneqq \varphi(\theta) - \varphi(\theta-1) = (-1)^\theta(2\theta - 1)$. Since N is even, $\varphi(-N) = -N$, and $\varphi(0) = 0$, so

$$N(-1)^{x_b} = -N^2 + \sum_S \left(\sum_{\theta=1-N}^{N+|S|m} d_\theta \,\big[\,X_S + t_S \ge \theta\,\big] - \sum_{\theta=1}^{|S|m} d_\theta \,\big[\,t_S \ge \theta\,\big] \right).$$

The bottom gates are the indicators $\big[\,X_S + t_S \ge \theta\,\big]$ and $\big[\,t_S \ge \theta\,\big]$, threshold gates by (1). The top gate outputs 1 exactly when the double sum is less than N^2, that is, when $x_b = 1$. The number of bottom gates is $\sum_S (2N + 2|S|m) = 2N^2 + 2m\sum_S |S| = 2N^2 + kmN$, where $\sum_S |S| = kN/2$ because each block $i \in \{1, \dots, k\}$ lies in half of the N sets S. In the balanced case $km = N$ and $n = 2N$, so this is $3N^2 = \frac{3}{4}n^2$. The top weights are $\pm d_\theta$. Since $-N < \theta \le N + km = n$, their absolute value is at most $2n - 1$. □

Since the bottom weights are small integers, the circuit is even a threshold of majority gates ($\mathsf{THR} \circ \mathsf{MAJ}$, by repeating inputs). For comparison, Kane and Williams also give a $\mathsf{MOD}_3 \circ \mathsf{MOD}_2$ circuit of size $O(n^2)$ [KW16, Thm. 1.2], in a different basis.

A sharper score: the 5/16 bound

Theorem 5. *Put $\alpha = \lfloor (m-1)/2 \rfloor$. There is an exact depth-two threshold circuit for A with at most*

$$\frac{N(N-1)}{2} + \frac{kN(m+\alpha)}{2} + 2(N-1) + 2\alpha k + 3km + \left\lceil \frac{km}{2} \right\rceil \tag{2}$$

bottom gates and one top gate. Its weights are integers of magnitude polynomial in n. In the balanced case its total size is

$$\frac{5}{16}n^2 - \frac{n}{4}\log_2(n/2) + O(n).$$

Idea of the proof. Since the top gate only tests the sign of $N(-1)^{x_b}$, any score that is at least N when $x_b = 0$ and at most $-N$ when $x_b = 1$ serves. With $Y_S = \sum_{c:\, S\cdot c \text{ odd}} x_c$ we have $X_S = 4Y_S - 2\sum_c x_c$ for $S \neq \varnothing$, so replacing X_S by $4Y_S$ in Lemma 1 changes the sum only by an affine term and a multiple of $[\, b = 0\,]$, and $u_S = 2Y_S + t_S$ reduces the width of the table-dependent part of the range from $2N$ to N. Further, $4Y_S(-1)^{t_S}$ equals $-8\lceil u_S/2 \rceil^2$ up to a polynomial in Y_S, t_S and $\lceil t_S/2 \rceil$, and the square changes only at odd u_S, so its staircase needs only odd thresholds. Summed over S, the polynomials collapse into $O(n)$ gates, up to errors at $b = 0$ and at the k addresses e_i of weight one. A term $-Cx_0[\, b = 0\,]$ with C large, built from the forms t_S with x_0 attached by a larger weight, repairs $b = 0$ at about $kmN/4$ gates. At $b = e_i$ the error is $-N|a_i|$ with $|a_i| \leq 2\alpha + 1$; a shift of the staircase coefficients handles $x_{e_i} = 0$, and weight $1 + \alpha$ on x_{e_i} inside Y_S handles $x_{e_i} = 1$. The full proof is in Appendix A.

The sharper construction remains sign-coherent: every bottom gate sees a block through its sum with a nonnegative coefficient. Mixed-sign weights are therefore not needed for this bound. Section 6 shows that improving the quadratic exponent requires gates that are not sign-coherent.

4 Agreeing on most inputs with fewer gates

Under uniform inputs every form $X_S \pm t_S$ concentrates in a window of width $O(\sqrt{n})$, so each staircase can be cut to its window. The question is how the N cut-off errors add up. They are not independent, but with a symmetric truncation they are orthogonal.

Lemma 6 (Odd error functions are orthogonal)**.** *Let the table be uniformly random. For distinct sets $S, T \subseteq \{1, \ldots, k\}$ and functions $g, h : \mathbb{Z} \to \mathbb{R}$ with g an odd function, that is, $g(-y) = -g(y)$ for all y, we have $\mathbb{E}[g(X_S)\, h(X_T)] = 0$.*

Proof. Since $S \neq T$, pick $u \in \{0,1\}^k$ with $(S \triangle T) \cdot u$ odd, and map the table x to x' with $(-1)^{x'_c} = (-1)^{T\cdot u}(-1)^{x_{c\oplus u}}$. This is a bijection of tables, so x' is uniform too, and

$$\begin{aligned} X_S(x') &= (-1)^{T\cdot u}\sum_c (-1)^{S\cdot(c\oplus u)+x_c} = (-1)^{S\cdot u + T\cdot u}\sum_c (-1)^{S\cdot c + x_c} \\ &= (-1)^{(S\triangle T)\cdot u} X_S(x) = -X_S(x), \end{aligned}$$

and in the same way $X_T(x') = X_T(x)$. Since x' is uniform,

$$\mathbb{E}_x[g(X_S(x))h(X_T(x))] = \mathbb{E}_{x'}[g(X_S(x'))h(X_T(x'))] = \mathbb{E}_x[g(-X_S(x))h(X_T(x))].$$

As g is an odd function, the right side is the negative of the left side, so both are zero. □

X_S and X_T are not independent; the lemma only uses this pairwise sign symmetry.

Theorem 7. *There are explicit depth-two threshold circuits for Andreev's function A with $m = N/k$ (the balanced case) as follows:*

1. *For $0 < \delta \le 1/2$, a circuit with $O\big(n^{3/2}\sqrt{\log(1/\delta)}\big)$ gates that agrees with A on at least a $1-\delta$ fraction of all inputs.*
2. *For $N^{-1/2} \le \varepsilon \le 1/8$, a circuit with $O\big(\varepsilon^2 n^{3/2}\sqrt{\log(1/\varepsilon)}\big)$ gates and $O\big(\varepsilon^2 n^{5/2}\sqrt{\log(1/\varepsilon)}\big)$ wires that ε-agrees with A, i.e. agrees on at least a $1/2+\varepsilon$ fraction.*

All these circuits have sign-coherent bottom gates, i.e. inside every address block the weights of a gate are all ≥ 0 or all ≤ 0.

Proof. A majority vote with fewer characters. Fix $0 \le r \le k$ and put $q = 2^r$. The proof of Lemma 1 with the sum restricted to $S \subseteq [r]$ gives

$$P_* \coloneqq 2\sum_{S\subseteq[r]} X_S(-1)^{t_S} = 2q \sum_{c:\, c_{[r]}=b_{[r]}} (-1)^{x_c}.$$

So P_* is a vote among $L = N/q$ table bits, one of them x_b. For $q = N$ we have $P_* = \pm 2N$ and output $\big[-P_* \ge 0\big]$; for $q < N$ we output the majority $\big[-P_* \ge 2q\big]$, with ties going to 0.

For $q < N$, fix the address input; the L voters are then independent uniform bits, and L is even. In the case where exactly $L/2$ of the other $L-1$ voters are 1, the vote follows x_b and is correct. Otherwise it does not depend on x_b, so it is correct with probability $1/2$. Hence its agreement is

$$\frac{1}{2} + \frac{1}{2}\cdot 2^{-(L-1)}\binom{L-1}{L/2} = \frac{1}{2} + 2^{-L-1}\binom{L}{L/2} \ge \frac{1}{2} + \frac{1}{2\sqrt{2L}}.$$

Symmetric truncation. For nonempty $S \subseteq [r]$ center the address count, $D_S = t_S - \lfloor m|S|/2 \rfloor$, and put $\eta_S = (-1)^{\lfloor m|S|/2 \rfloor}$. Lemma 2, applied with $t = D_S$ and $t = -D_S$, gives

$$\eta_S\big(\varphi(X_S + D_S) + \varphi(X_S - D_S)\big) = 2X_S(-1)^{t_S}.$$

For an integer radius $R \ge 1$, let P_R be obtained from $P_* = 2X_\varnothing + \sum_{S\ne\varnothing} \eta_S\big(\varphi(X_S+D_S)+\varphi(X_S - D_S)\big)$ by replacing every $\varphi(u)$ by its truncation $\varphi(u)\big[\,|u| \le R\,\big]$ (the affine term $2X_\varnothing$ stays). Each truncated staircase has at most $2R+2$ steps, so P_R uses at most $4(q-1)(R+1)$ bottom gates. Their address weights are $+1$ on the blocks of S in one form and -1 in the other, so they are sign-coherent.

Cancellation of cross terms. Let e_S be the term of S in P_R minus its term in P_*; the affine term cancels, so $P_R - P_* = \sum_{S\ne\varnothing} e_S$. Fix the address input. Then D_S and η_S are constants and

$$e_S = -\eta_S\big(\varphi(X_S+D_S)\big[\,|X_S+D_S| > R\,\big] + \varphi(X_S-D_S)\big[\,|X_S-D_S| > R\,\big]\big) =: g_S(X_S)$$

is a function of X_S alone. The function g_S is odd because $u \mapsto \varphi(u)\big[\,|u| > R\,\big]$ is an odd function: $\varphi(-u) = -\varphi(u)$ and the cutoff depends only on $|u|$. Replacing y by $-y$ sends $y + D_S$ to $-(y - D_S)$ and $y - D_S$ to $-(y + D_S)$, so the two summands are exchanged and negated, giving $g_S(-y) = -g_S(y)$. The table stays uniform when the address input is fixed, so by Lemma 6 with $g = g_S$ and $h = g_T$ we get $\mathbb{E}[e_S e_T] = 0$ for all $S \ne T$, first for each fixed address and then on average. Therefore

$$\mathbb{E}(P_R - P_*)^2 = \sum_{S\ne\varnothing} \mathbb{E}\, e_S^2 + \sum_{S\ne T} \mathbb{E}[e_S e_T] = \sum_{S\ne\varnothing} \mathbb{E}\, e_S^2.$$

Diagonal terms. Here the expectations are over both the table and the address bits. For $U = X_S \pm D_S$ we have $|\mathbb{E}U| \le 1/2$, and $U - \mathbb{E}U$ is a sum of N independent uniform signs and $m|S|$ independent uniform variables in $\{-1/2, 1/2\}$, so Hoeffding's inequality gives $\Pr\big[|U - \mathbb{E}U| \ge \lambda\big] \le 2\exp(-\lambda^2/(2v))$ with $v = N + km/4$. As $|U| > R$ implies $|U - \mathbb{E}U| > R/2$ and $U^2 \le 4(U - \mathbb{E}U)^2$, integrating this tail from $R/2$ gives $\mathbb{E}\big[U^2\big[\,|U| > R\,\big]\big] \le (2R^2 + 16v)e^{-R^2/(8v)} \le 32v\exp\big(-R^2/(16v)\big)$, and $e_S^2 \le 2\sum_{\pm} U^2\big[\,|U| > R\,\big]$ gives

$$\mathbb{E}(P_R - P_*)^2 = \sum_{S \ne \varnothing} \mathbb{E}\, e_S^2 \le 128(q-1)v\exp(-R^2/(16v)).$$

For $q < N$, the sum $P_*/(2q)$ contains an even number L of signs, so P_* is a multiple of $4q$ and lies at distance at least $2q$ from the threshold $-2q$. For $q = N$, the two scores $\pm 2N$ have the same margin from the threshold 0. Thus the truncated circuit differs from the vote only if $|P_R - P_*| \ge 2q$. By Markov's inequality this has probability at most $(32v/q)\exp(-R^2/(16v))$, which is at most δ for

$$R \coloneqq \left\lceil 4\sqrt{v \ln \frac{32v}{q\delta}} \right\rceil.$$

The two parameter choices. For Part 1 take $q = N$, so the vote is A itself; at balance $v = O(N)$ and $R = O(\sqrt{N\log(1/\delta)})$. For Part 2 take the least power of two $q \ge 32\varepsilon^2 N$ and $\delta = \varepsilon$. Then $32 \le q \le N/2$ because $N^{-1/2} \le \varepsilon \le 1/8$, and $q < 64\varepsilon^2 N$; the vote agrees on at least a $1/2 + \sqrt{q/(8N)} \ge 1/2 + 2\varepsilon$ fraction, and truncation loses at most ε. Every bottom gate has at most n input wires and one wire to the top gate, which also has at most n direct input wires. This gives the stated wire bound. □

5 Lower bounds

This section proves lower bounds on the gates and wires of an arbitrary depth-two threshold circuit that agrees with Andreev's function on at least a $1/2 + \varepsilon$ fraction of inputs; $\varepsilon = 1/2$ means exact computation. For gates, the bound matches Theorem 7 up to polylogarithmic factors. The argument is the random restriction method of Kane and Williams [KW16] with two changes taken from Section 2: a sharper count of restricted circuits saves a logarithm, and a union bound over all survivor counts improves ε^3 to ε^2.

Theorem 8. *Let C be a depth-two threshold circuit with s gates and w wires that ε-agrees with Andreev's function, where $\varepsilon \ge 3(k+2)/\sqrt{N}$. Then, in the balanced case ($km = N$),*

$$s = \Omega(\varepsilon^2 n^{3/2}/\log^2 n) \quad \textit{and} \quad w = \Omega(\varepsilon^2 n^{5/2}/\log^{5/2} n).$$

Taking $\varepsilon = \frac{1}{2}$, allowed when $N \ge 36(k+2)^2$, gives the worst-case bounds.

Plan. We follow Kane and Williams [KW16]. Fix all but one bit per block at random. The restricted target function is a lookup in a *random* table indexed by the k remaining bits, and only a few bottom gates still depend on those bits (Lemma 9). There are few such restricted circuits (Lemma 11), and with high probability none correlates well with a random table (Lemma 12). Averaging over all restrictions proves the theorem.

Restrictions. The lemmas below hold for all $k, m \geq 1$, not only in the balanced case. Throughout, C is a depth-two threshold circuit of size s with w wires. A *restriction* consists of a choice σ of one *live* position in every block, and a *fixing*: values for all other input bits, table included. For a uniformly random restriction, the live position in each block is chosen uniformly and independently, and all fixed bits are independent uniform bits, independent of these choices. Write $u \in \{0,1\}^k$ for the live bits and o_i for the parity of the $m-1$ fixed bits of block i (the *offset*). Then the address is $b = u \oplus o$, and the restricted function is $u \mapsto x_{u \oplus o}$. A bottom gate *survives* if it still depends on u. Let U be the number of surviving gates.

Lemma 9 (Few survivors)**.** *For a uniformly random restriction, $\mathbb{E}[U] \leq 4s\sqrt{n}/m$. The expected number of bottom gates that depend on at least two live bits is at most $4w/m^{3/2}$. Both bounds also hold for every fixed table, averaging over the live positions and the fixed address bits.*

These are the random restriction lemmas of Kane and Williams [KW16, Lemmas 1.1 and 3.1], with explicit constants. Appendix B proves both bounds; the proof includes small repairs to their argument (ties among weights, zero weights, and a strict versus non-strict inequality). Chen, Tal and Wang give an alternative proof of the first lemma, up to constants [CTW26, App. A.3]; Kabanets, Kane and Lu [KKL17, Lemma 6.1] and Chen, Santhanam and Srinivasan [CSS18, Lemma 9.1] prove analogues with polylogarithmic and $n^{o(1)}$ losses.

Lemma 10 (Roychowdhury, Siu, Orlitsky [RSO94]; see [KW16, Thm. 2.3])**.** *For any fixed Boolean functions $f_1, \ldots, f_q : \Omega \to \{0,1\}$ on a finite set Ω, at most $(|\Omega|+1)^{q+1}$ functions on Ω have the form $[\sum_j \omega_j f_j \geq \tau]$.*

Lemma 11 (Counting restricted circuits)**.** *Fix C and σ. Every fixing turns C into a function of the live bits $u \in \{0,1\}^k$. For an integer $q \geq 0$, let $\mathcal{H}_q$ be the set of functions that arise in this way from fixings with at most q survivors. Then*

$$|\mathcal{H}_q| \leq \left((2^k+1)(s+1)\right)^q (2^k+1)^{q+k+1}.$$

Proof. Non-surviving bottom gates are constants, and are absorbed into the top threshold. A surviving gate is $[\omega \cdot u \geq \tau]$, where ω is its original weight vector restricted to the live positions; only τ depends on the fixing. For a fixed ω, these functions are nested as τ grows, so there are at most 2^k+1 of them. Pad the list of survivors with constant functions to length exactly q. Then each entry of the list is one of at most $(2^k+1)(s+1)$ functions: choose one of the at most s gates or "constant", then the threshold. The top gate is a threshold function of the q entries and the k live bits. By Lemma 10 with $|\Omega| = 2^k$, there are at most $(2^k+1)^{q+k+1}$ such functions for each list. □

Kane and Williams [KW16], and similarly Kabanets, Kane and Lu [KKL17] and Chen, Santhanam and Srinivasan [CSS18], count all threshold functions of the live bits at this step, which costs about k^2 bits per survivor instead of $O(k + \log s)$.

Lemma 12 (Agreement bound for all survivor counts at once)**.** *Let $\alpha = \sqrt{N(k+2+\ln(s+1))}$ and $\beta = (k+2)\sqrt{N}$. Then C agrees with Andreev's function on at most a*

$$\tfrac{1}{2} + \frac{\alpha\sqrt{\mathbb{E}[U]} + \beta}{N} + 2^{-(k+6)}$$

fraction of inputs.

Proof. Fix σ and the fixed address bits; the table remains uniformly random. For one function f of u and $\Delta \geq 0$, Hoeffding's inequality gives

$$\Pr[f \text{ agrees with } u \mapsto x_{u\oplus o} \text{ on} \geq N/2 + \Delta \text{ points}] \leq e^{-2\Delta^2/N}.$$

Put $\Delta_q = \alpha\sqrt{q} + \beta$. Call the table *bad* if, for some integer $0 \leq q < N$, a function $f \in \mathcal{H}_q$ agrees with $u \mapsto x_{u\oplus o}$ on more than $N/2 + \Delta_q$ points. Using $\Delta_q^2 \geq \alpha^2 q + \beta^2$, $\ln(2^k + 1) \leq k + 1$ and Lemma 11,

$$\begin{aligned}\Pr[\text{bad}] &\leq \sum_{q\geq 0} |\mathcal{H}_q|\, e^{-2\Delta_q^2/N} \leq \sum_{q\geq 0} e^{(2q+k+1)(k+1)+q\ln(s+1)-2q(k+2+\ln(s+1))-2(k+2)^2} \\ &= \sum_{q\geq 0} e^{-2q-q\ln(s+1)+(k+1)^2-2(k+2)^2} \leq \sum_{q\geq 0} e^{-2q-(k^2+6k+7)},\end{aligned}$$

which is at most $1.2\, e^{-(k^2+6k+7)} \leq 2^{-(k+6)}$.

For a good table with $U < N$ survivors, the agreement on the N points u is at most $N/2+\alpha\sqrt{U}+\beta$. This also holds for $U \geq N$, because then $\alpha\sqrt{U} \geq N > N/2$. Averaging over all restrictions and using $\mathbb{E}\sqrt{U} \leq \sqrt{\mathbb{E}U}$ (Cauchy–Schwarz) proves the lemma, since the average agreement over all restrictions equals the agreement over all inputs. □

Kane and Williams [KW16], and similarly the later work [KKL17, CSS18], instead bound U by Markov's inequality and then use a single survivor count. That loses a factor of ε; the uniform bound with Cauchy–Schwarz does not.

Proof of Theorem 8. We prove more generally, for all $k, m \geq 1$ and $\varepsilon \geq 3(k+2)/\sqrt{N}$, that

$$m\, N\, \varepsilon^2 \leq 36\,(k + 2 + \ln(s+1))\, s\sqrt{n} \quad \text{and} \quad m^{3/2} N \varepsilon^2 \leq 36\,(k + 2 + \ln(w+1))\, w, \tag{3}$$

and that for $m \leq N$ the factors $k + 2 + \ln(s + 1)$ and $k + 2 + \ln(w + 1)$ can both be replaced by $4(k + 2)$:

$$mN\varepsilon^2 \leq 144(k+2)\, s\sqrt{n} \quad \text{and} \quad m^{3/2} N \varepsilon^2 \leq 144(k+2)\, w. \tag{4}$$

In the balanced case $m = N/k \leq N$ and $n = 2N$, so (4) reads

$$s \geq \frac{\varepsilon^2 N^{3/2}}{144\sqrt{2}\, k(k+2)} \quad \text{and} \quad w \geq \frac{\varepsilon^2 N^{5/2}}{144\, k^{3/2}(k+2)},$$

which is the theorem, since $N = n/2$ and $k = \log_2(n/2)$.

By hypothesis, $\beta/N = (k+2)/\sqrt{N} \leq \varepsilon/3$ and $2^{-(k+6)} \leq \varepsilon/192$. So Lemma 12 gives $\varepsilon/3 \leq \alpha\sqrt{\mathbb{E}[U]}/N$. Squaring and inserting Lemma 9,

$$\varepsilon^2 N^2 \leq 9\alpha^2\, \mathbb{E}[U] \leq 9N\big(k + 2 + \ln(s+1)\big) \cdot \frac{4s\sqrt{n}}{m},$$

which is the first bound of (3).

For wires, count as survivors only the gates that depend on at least two live bits. A gate depending on at most one live bit computes that bit, its negation, or a constant, so it can be absorbed into the top gate, which reads the live bits anyway. Every remaining survivor has at least two input wires, so there are at most w candidate gates, and Lemmas 11 and 12 hold with s replaced by w. Now use the second bound of Lemma 9.

For (4), let $m \leq N$. If $\ln(s + 1) \leq 3(k + 2)$, then $k + 2 + \ln(s + 1) \leq 4(k + 2)$. Otherwise $s \geq e^{3(k+2)} - 1 > N^{5/2} \geq mN$, so the first bound holds trivially. The same argument works for w, since $N^{5/2} \geq m^{3/2}N$. □

Remark (Limits of the method). Theorem 8 is within one logarithm of what survivor counting can give. After a restriction with one live bit per block, the function $u \mapsto x_{u \oplus o}$ has a depth-two circuit with N bottom gates, and gates such as $[\, t_S \geq \theta \,]$ survive with probability $\Theta(k/\sqrt{N})$; Kane and Williams already observe that their restriction lemma is tight for majority gates [KW16, Section 3]. So counting the survivors of a single restriction cannot give more than about $n^{3/2}/\log n$ gates.

6 A near-quadratic lower bound for sign-coherent circuits

The upper bound circuit of Theorem 5 uses only bottom gates of a special kind. This section shows that with such gates it cannot be improved by more than a polylogarithmic factor.

Sign-coherence. A bottom gate is *sign-coherent* if, inside every address block, its weights are either all ≥ 0 or all ≤ 0; its table weights are arbitrary. A depth-two threshold circuit is *sign-coherent* if all its bottom gates are; its top gate is arbitrary. Every gate whose address weights are constant within each block is sign-coherent, as is every gate with nonnegative address weights. All bottom gates of the circuit in Theorem 5 are of this kind.

Theorem 13. *Let C be a sign-coherent depth-two threshold circuit for A with G bottom gates. Put $D_0 = \max\{0, N - k - 1 - \log_2(G+1)\}/(k + \log_2(G+1))$. Then*

$$G \;\geq\; \frac{m^k}{m^k - (m-1)^k}\, D_0 \;\geq\; \frac{m}{k}\, D_0.$$

Consequently $G \geq \min\{N^2,\ m(N-3k-1)/(3k^2)\}$, and in the balanced case

$$G \;\geq\; \frac{N(N-3k-1)}{3k^3} \;=\; \Omega\!\left(\frac{n^2}{\log^3 n}\right).$$

So among sign-coherent circuits the circuit of Theorem 5 is optimal up to a polylogarithmic factor. The earlier circuit of Kane and Williams [KW16, Thm. 1.2] is sign-coherent as well. A polynomial-exponent improvement below quadratic therefore requires gates whose weights have mixed signs inside a block.

Plan. We walk through all unit cubes of the grid of block counts. On each cube, the circuit reveals the whole table, so many gates must be non-constant there (Lemma 14). A sign-coherent gate is non-constant on at most a k/m fraction of the cubes (Lemma 15). Comparing the two counts gives the theorem.

The cube restriction. Fix an order of the positions inside every block. For a vector $s \in \{0, \ldots, m\}^k$ let $a(s)$ be the address part in which block i has ones exactly at its first s_i positions. A *cube* is a vector $q \in \{0, \ldots, m-1\}^k$; there are m^k cubes. For $z \in \{0,1\}^k$ the input $(x, a(q+z))$ differs from $(x, a(q))$ exactly in the live positions $q_i + 1$ of the blocks i with $z_i = 1$, and its address is $(q \bmod 2) \oplus z$. Hence, for every cube q,

$$z \mapsto A\big(x, a(q+z)\big) = x_{(q \bmod 2) \oplus z}$$

determines the whole table x. A bottom gate *splits* the cube q for the table x if it is non-constant in z on the inputs $(x, a(q+z))$. Let $d_q(x)$ be the number of bottom gates that split q for x.

Lemma 14 (Many gates split each cube)**.** *Let C be any depth-two threshold circuit for A with G bottom gates, and let x be uniformly random. Then $\mathbb{E}_x[d_q(x)] \geq D_0$ for every cube q.*

Proof. Fix q and x, and consider bottom gate j as a function of z alone. It computes $[\,\lambda_j \cdot z \geq \tau_j(x)\,]$, where $\lambda_j \in \mathbb{R}^k$ lists its fixed weights on the live positions, and $\tau_j(x)$ absorbs the contributions of the table bits and fixed address bits. Let I be the set of gates that split q for this x, and write $g_j(z)$ for their restricted functions. Absorb the constant gates, the table bits and the fixed address bits into the top threshold $\kappa(x)$. The top gate then computes $[\,\sum_{j\in I} v_j g_j(z) + \mu \cdot z \geq \kappa(x)\,]$, where v_j are its fixed weights on the bottom gates and μ lists its fixed weights on the live positions. So among the weights and thresholds, only τ_j and κ depend on x.

Therefore the function $z \mapsto x_{(q \bmod 2)\oplus z}$, and with it x, is determined by: the set I of splitting gates; for each $j \in I$ the non-constant function $[\,\lambda_j \cdot z \geq \tau_j(x)\,]$, one of at most $2^k - 1$ (only the position of $\tau_j(x)$ among the $\leq 2^k$ values $\lambda_j \cdot z$ matters, and the two constant functions are excluded); and the position of $\kappa(x)$. After the restricted bottom-gate functions have been specified, the top gate's score is fixed at each of the 2^k inputs z, so at most $2^k + 1$ positions of $\kappa(x)$ matter. Now let x be uniformly random, and write H for Shannon entropy in bits. Encoding $|I| = d_q(x)$ first, the entropy chain rule gives

$$\begin{aligned} N = H(x) &\leq H(d_q(x)) + \mathbb{E}_x\left[\log_2\left(\binom{G}{d_q(x)}(2^k-1)^{d_q(x)}(2^k+1)\right)\right] \\ &\leq \log_2(G+1) + \mathbb{E}_x[d_q(x)]\,(\log_2(G+1)+k) + k + 1, \end{aligned}$$

which rearranges to the claim; a negative numerator is replaced by the trivial lower bound zero. □

Lemma 15 (A coherent gate splits few cubes)**.** *For every table x, a sign-coherent gate splits at most $m^k - (m-1)^k$ cubes.*

Proof. Fix the table x and a sign-coherent gate with threshold θ. Write $w_{i\ell}$ for its weight on position ℓ of block i. On the input $(x, a(s))$, its score is

$$F(s) = c(x) + \sum_{i=1}^{k}\sum_{\ell=1}^{s_i} w_{i\ell},$$

where $c(x)$ is the contribution of the table bits. Let $\varepsilon_i = +1$ if all weights in block i are nonnegative, and $\varepsilon_i = -1$ otherwise. By sign-coherence, moving s_i one step in direction ε_i never decreases F.

For a cube q, define the two opposite vertices

$$q^- = q + (\mathbf{1} - \varepsilon)/2, \qquad q^+ = q + (\mathbf{1} + \varepsilon)/2.$$

These minimize and maximize F on the cube, respectively. Thus the gate splits q if and only if $F(q^-) < \theta \leq F(q^+)$.

Let $Q = \{0, \ldots, m-1\}^k$ be the set of cubes. Partition Q into maximal *lines* q, $q+\varepsilon$, $q+2\varepsilon, \ldots$ contained in Q. For consecutive cubes, the identity

$$(q + \varepsilon)^- = q^+$$

shows that the largest score on one cube equals the smallest score on the next. Hence their half-open intervals $(F(q^-), F(q^+)]$ are disjoint along each line, and the fixed threshold θ can split at most one cube per line.

Each line has exactly one starting cube: a $q \in Q$ with $q - \varepsilon \notin Q$. To count these, note that $q - \varepsilon \in Q$ requires $q_i \in \{1, \ldots, m-1\}$ when $\varepsilon_i = +1$, and $q_i \in \{0, \ldots, m-2\}$ when $\varepsilon_i = -1$. There are $m - 1$ choices in each coordinate, hence $(m-1)^k$ cubes with a predecessor. The remaining $m^k - (m-1)^k$ cubes start lines, proving the bound. □

Proof of Theorem 13. Sum Lemma 14 over all cubes and compare with Lemma 15: $m^k D_0 \le \sum_q \mathbb{E}_x[d_q(x)] \le G\,(m^k - (m-1)^k)$. Then use $1-(1-1/m)^k \le k/m$. If $G \ge N^2$ there is nothing to prove. Otherwise $\log_2(G+1) \le 2k$ and $D_0 \ge (N-3k-1)/(3k)$. In the balanced case $m = N/k$. □

Remark. (1) Lemma 15 is tight for all k, m. The fact that a unate function changes at most once along a chain compatible with its orientation is classical [IPS97, Lemma 2.1]; to the best of our knowledge, the cube restriction and the resulting near-quadratic lower bound are new. (2) The explicit balanced bound in Theorem 13 first exceeds $n^{3/2}$ at $k = 64$ among the admissible powers of two. The method (Lemma 14 plus a uniform bound as in Lemma 15) cannot give more than $O(n^2/\log n)$. (3) Lemma 14 holds for all circuits. For general gates, the limit of survivor counting noted after Theorem 8 does not depend on the choice of restrictions: under every distribution of restrictions that leaves one live bit per block, a gate with random ± 1 weights is non-constant with probability at least $1/(2\sqrt{2km})$, so such arguments cannot exceed $O(n^{3/2})$. Gates with mixed signs inside a block are exactly where the gap between $n^{3/2}$ and n^2 lies.

7 Open questions

1. *Mixed signs.* Is there an exact circuit with $n^{2-\Omega(1)}$ gates? By Theorem 13 it must use bottom gates whose weights have mixed signs inside a block. Conversely, can the exact lower bound beat $n^{3/2}$? Such a proof must handle gates whose address weights nearly cancel, since those defeat both survival counting and the cube restriction. Rectangle methods (discrepancy, sign rank) do not reach $n^{2-o(1)}$ either: for every split of the inputs between two players, Andreev's function has a protocol with $O(k)$ bits of communication, and sign rank is too weak against $\mathsf{THR} \circ \mathsf{THR}$ in general [CM17].

2. *Parity.* Fixing the table to $x_c = y \oplus c_1 \oplus \cdots \oplus c_k$ turns a circuit for A into one for parity of $km+1$ bits. With sign-coherent gates parity needs about half as many gates as bits; with general gates a linear depth-two lower bound for parity is open and tied to slicing all edges of the hypercube with hyperplanes. Does a direct-sum version, about N times the parity cost, hold for Andreev's function?

3. *Wires.* The exact wire bounds are $\Omega(n^{5/2}/\log^{5/2} n)$ and $O(n^3)$. Can either exponent move?

4. *From approximation to exactness.* The gate complexity is about $n^{3/2}$ for constant error and between $n^{3/2}$ and n^2 for error 0. How does it depend on the error δ between $2^{-\Theta(n)}$ and a constant?

Formalization. The Lean 4 proofs are in the `cslib-frontier` repository at `https://github.com/cslib-community/cslib-frontier`. We refer to commit
`89ed4de0e54b4953fcc15b8c5d33fad4a344ea84`
on the `main` branch. The files are in `WeightedThreshold/` under `CslibFrontier/Computability/Circuit/Boolean/`. The main files are:

- `Andreev.lean` and `AndreevSharp.lean` for the exact constructions (Proposition 4 and Theorem 5);
- `AndreevOddErrors.lean` and `AndreevTruncatedCircuit.lean` for Lemma 6 and Theorem 7;
- `AndreevLowerBound.lean` for Theorem 8;

- `AndreevSignCoherent.lean` for Theorem 13 and Lemmas 14–15.

All numbered results are covered by these files and their dependencies. The formalization proves explicit bounds in the parameters; the asymptotic consequences are derived in the text. The remarks are not formalized.

Acknowledgements. The results, the formalization and this note were developed with AI assistance.

A Proof of Theorem 5

Proof. Forms used only in this proof. Write $s_i = |a_i|$, $\sigma = \sum_i s_i$, and let e_i be the address with only coordinate i equal to one. Put

$$z \coloneqq \sum_{c\neq 0} x_c, \quad E \coloneqq \sum_i x_{e_i}, \quad w \coloneqq 2z + 2\alpha E + \sigma, \quad R \coloneqq 2(N-1) + 2\alpha k + km, \quad C \coloneqq N(R+4).$$

Thus $0 \le w \le R$. For each nonempty $S \subseteq [k]$, define

$$Y_S \coloneqq \sum_{c: S\cdot c \text{ odd}} x_c, \quad W_S \coloneqq Y_S + \alpha \sum_{i\in S} x_{e_i}, \quad \kappa_S \coloneqq -4(2\alpha+1)(k-2|S|),$$

and let $r_S(u) = -8\lceil u/2\rceil^2 + \kappa_S\lceil u/2\rceil$. All of W_S, t_S, w, s_i are affine forms in the original inputs.

The score. Consider

$$\begin{aligned} P \coloneqq 2N + \sum_{S\neq\varnothing} r_S(2W_S + t_S) + \frac{N}{2}\left(w^2 + \sum_i s_i^2\right) \\ + 4Nz - 2N(\alpha^2+1)E - 2\alpha N\sigma + 2N(1+\alpha)\sum_i x_{e_i}s_i - Cx_0[\,b=0\,]. \end{aligned} \tag{5}$$

We first verify its sign, then realize its non-affine terms with bottom gates. The following identities follow by expanding the squares and summing characters; all sums over S are over nonempty sets:

$$\begin{aligned} \sum_S 4W_S(-1)^{t_S} &= 2N[\,b=0\,]z - 2N[\,b\neq 0\,]x_b + 2\alpha N\sum_i x_{e_i}([\,b=0\,] - [\,b=e_i\,]), \\ 8\sum_S (W_S^2 + W_S t_S) &= \frac{N}{2}(w^2-\sigma^2) + 2Nz + (4\alpha+2\alpha^2)NE + 2N(1+\alpha)\sum_i x_{e_i}s_i, \\ 8\sum_S \lceil t_S/2\rceil^2 &= \frac{N}{2}\left(\sigma^2 + \sum_i s_i^2\right) + N(\sigma+1)[\,b\neq 0\,] + N\sum_i s_i[\,b=e_i\,]. \end{aligned}$$

For example, the first is character orthogonality applied to Y_S and to $[\,i\in S\,]$; the other two use the same identity and $\lceil t/2\rceil = (t + [\,t \text{ odd}\,])/2$. Also

$$4W(-1)^t = 4W - 8\lceil (2W+t)/2\rceil^2 + 8W^2 + 8Wt + 8\lceil t/2\rceil^2.$$

The coefficient shifts κ_S supply the address correction because

$$\sum_S \kappa_S(\lceil t_S/2\rceil - t_S/2) = N(2\alpha+1)\sum_i [\,b=e_i\,].$$

Substitution in (5) now gives the simple closed form

$$P = \begin{cases} N(w+2) - Cx_0, & b = 0, \\ 2N(1+\alpha)(1-x_{e_i}) - Ns_i, & b = e_i, \\ N(1-2x_b), & \text{otherwise.} \end{cases} \tag{6}$$

At $b = 0$, use $0 \le w \le R$ and $C = N(R+4)$. At $b = e_i$, the odd integer s_i lies between 1 and $2\alpha+1$. In every case $P \ge N$ when $x_b = 0$, and $P \le -N$ when $x_b = 1$. The top gate outputs $[\,-P \ge 0\,]$.

Raw gates and their count. The staircase for $r_S(2W_S + t_S)$ uses only odd thresholds

$$[\,2W_S + t_S \ge \theta\,], \qquad 1 \le \theta \le N + 2\alpha|S| + |S|m,$$

with coefficient $-8\theta + \kappa_S$. The two square terms use $R + km$ gates by Lemma 3. For $B = n + 2$, each product $x_{e_i}s_i$ uses m gates $[\,s_i + Bx_{e_i} \ge B + \ell\,]$, $1 \le \ell \le m$. Finally, the identity

$$x_0[\,b = 0\,] = x_0 - x_0\sigma + \frac{4}{N}\sum_{S \ne \varnothing} x_0 \lfloor t_S/2 \rfloor$$

implements the last correction with km gates $[\,\sigma + Bx_0 \ge B + \ell\,]$ and $\sum_S \lfloor |S|m/2 \rfloor$ gates $[\,t_S + Bx_0 \ge B + 2j\,]$. Its identity follows from $\sum_S [\,t_S \text{ odd}\,] = (N/2)[\,b \ne 0\,]$. The even-level σ gates coincide with the $S = [k]$ gates, saving $\lfloor km/2 \rfloor$ gates. Everything else in P is an affine top term.

Since $\sum_{S \ne \varnothing} |S| = kN/2$, the gate counts for the r_S staircases and the last correction sum to

$$\sum_{S \ne \varnothing} \left(\left\lceil \frac{N + 2\alpha|S| + |S|m}{2} \right\rceil + \left\lfloor \frac{|S|m}{2} \right\rfloor \right) = \frac{N(N-1)}{2} + \frac{\alpha k N}{2} + \frac{kmN}{2}.$$

Adding the remaining gates gives (2). Vanishing coefficients or further coincidences can only improve it. In the balanced case m is even, so $\alpha = m/2 - 1$; substitution gives the stated asymptotic bound. Bottom coefficients and thresholds are $O(n)$; the displayed expansion has top coefficients bounded by $O(n^3)$. □

B Proof of Lemma 9

Lemma 16 (Erdős [Erd45])**.** *Let $R \ge 1$ be an integer, let $z_1, \dots, z_R$ be independent uniform bits, and let real weights $\omega_1, \dots, \omega_R$ satisfy $|\omega_j| \ge \mu > 0$, and let I be a half-open interval of length μ. Then* $\Pr[\sum_j \omega_j z_j \in I] \le \binom{R}{\lfloor R/2 \rfloor} 2^{-R} \le 1/\sqrt{R}$.

Fix one bottom gate $[\,\sum_j \omega_j z_j \ge \tau\,]$, where j runs over all input positions. Let Λ be the set of address positions with $\omega_j \ne 0$, so $|\Lambda|$ is at most the number of the gate's input wires. For $j \in \Lambda$ let $\rho(j)$ be the number of $j' \in \Lambda$ with $|\omega_{j'}| \ge |\omega_j|$, including j itself. Fix the table bits; they only shift the threshold.

Lemma 17. *Fix the live positions. Over the random fixed address bits, the gate is non-constant on the live bits with probability at most $2\sum_j \rho(j)^{-1/2}$, the sum over the live positions $j \in \Lambda$.*

Proof. Let Φ be the contribution of the fixed bits. The restricted gate is $[\,\sum_{j \text{ live}} \omega_j u_j \ge \tau - \Phi\,]$, and it is non-constant if and only if $\tau - \Phi$ lies in a half-open interval of length $\sum_{j \text{ live}} |\omega_j|$. Split this interval into consecutive half-open pieces of lengths $|\omega_j|$, one for each live $j \in \Lambda$.

For a live $j \in \Lambda$, let R_j be the number of *fixed* positions $j' \in \Lambda$ with $|\omega_{j'}| \geq |\omega_j|$. If $R_j \geq \rho(j)/4$ for every live $j \in \Lambda$, then Lemma 16, applied to these R_j positions and conditioned on all other fixed bits, bounds the probability of piece j by $R_j^{-1/2} \leq 2\rho(j)^{-1/2}$, and summing over the pieces proves the lemma.

Otherwise some live j has $R_j < \rho(j)/4$, so more than $3\rho(j)/4$ live positions $j' \in \Lambda$ have $|\omega_{j'}| \geq |\omega_j|$, and each of them has $\rho(j') \leq \rho(j)$. Then the claimed bound is at least $\frac{3}{4}\rho(j) \cdot 2\rho(j)^{-1/2} \geq \frac{3}{2}$, which exceeds any probability. □

Proof of Lemma 9. Each position of a block is live with probability $1/m$. Sorting the values $\rho(j)$, the t-th smallest is at least t. So the gate survives with probability at most

$$\frac{2}{m} \sum_{j \in \Lambda} \rho(j)^{-1/2} \leq \frac{2}{m} \sum_{t=1}^{|\Lambda|} t^{-1/2} \leq \frac{4\sqrt{|\Lambda|}}{m} \leq \frac{4\sqrt{n}}{m}.$$

Summing over the at most s bottom gates gives the first bound.

A gate that depends on at least two live bits is non-constant and has at least two live positions in Λ. Positions in different blocks are live independently, so for $j \in \Lambda$,

$$\Pr[\, j \text{ and another position of } \Lambda \text{ are live}\,] \leq \min\Big\{\frac{1}{m}, \frac{|\Lambda|}{m^2}\Big\} \leq \frac{\sqrt{|\Lambda|}}{m^{3/2}}.$$

By Lemma 17, the gate depends on at least two live bits with probability at most

$$2 \sum_{j \in \Lambda} \frac{\sqrt{|\Lambda|}}{m^{3/2}}\, \rho(j)^{-1/2} \leq \frac{2\sqrt{|\Lambda|}}{m^{3/2}} \cdot 2\sqrt{|\Lambda|} = \frac{4|\Lambda|}{m^{3/2}}.$$

Summing over the bottom gates, and using that $\sum |\Lambda|$ is at most the number of wires w, gives the second bound. □

The formal proofs are the Lean theorems `card_not_forcedToConst_ltfEval_le` (Lemma 17, there with all non-live positions random), `sum_circuitUnforcedConstCount_blockLiveSet_le` and `sum_circuitUnforcedAtMostOneCount_blockLiveSet_le`.

References

[ACW16] J. Alman, T. M. Chan, and R. Williams. *Polynomial Representations of Threshold Functions and Algorithmic Applications.* FOCS 2016, pages 467–476.

[And87] A. E. Andreev. *On a method for obtaining more than quadratic effective lower bounds for the complexity of π-schemes.* Moscow Univ. Math. Bull. 42(1):63–66, 1987.

[CM17] A. Chattopadhyay and N. S. Mande. *Weights at the Bottom Matter When the Top is Heavy.* ECCC TR17-083, 2017.

[CSS18] R. Chen, R. Santhanam, and S. Srinivasan. *Average-Case Lower Bounds and Satisfiability Algorithms for Small Threshold Circuits.* Theory of Computing 14(9):1–55, 2018.

[CTW26] L. Chen, A. Tal, and Y. Wang. *Superquadratic Lower Bounds for Depth-2 Linear Threshold Circuits.* STOC 2026, pages 874–885. ECCC TR26-039.

[Erd45] P. Erdős. *On a lemma of Littlewood and Offord*. Bull. Amer. Math. Soc. 51:898–902, 1945.

[HMPST93] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy, and G. Turán. *Threshold circuits of bounded depth*. J. Comput. Syst. Sci. 46(2):129–154, 1993.

[Hås98] J. Håstad. *The shrinkage exponent of de Morgan formulas is* 2. SIAM J. Comput. 27(1):48–64, 1998.

[IPS97] R. Impagliazzo, R. Paturi, and M. Saks. *Size–depth tradeoffs for threshold circuits*. SIAM J. Comput. 26(3):693–707, 1997.

[KKL17] V. Kabanets, D. M. Kane, and Z. Lu. *A Polynomial Restriction Lemma with Applications*. STOC 2017, pages 615–628. ECCC TR17-026.

[KW16] D. M. Kane and R. Williams. *Super-Linear Gate and Super-Quadratic Wire Lower Bounds for Depth-Two and Depth-Three Threshold Circuits*. STOC 2016, pages 633–643. arXiv:1511.07860v1.

[Nag26] D. Nag. *Almost-Everywhere Near-Cubic Wire Lower Bounds for* $\mathsf{SYM}\circ\mathsf{THR}$ *and* $\mathsf{THR}\circ\mathsf{THR}$. Preprint, ECCC TR26-167, 2026.

[Nis93] N. Nisan. *The communication complexity of threshold gates*. In Combinatorics, Paul Erdős is Eighty, Vol. 1, pages 301–315, 1993.

[RSO94] V. P. Roychowdhury, K.-Y. Siu, and A. Orlitsky. *Neural models and spectral methods*. In Theoretical Advances in Neural Computation and Learning, pages 3–36, 1994.

[Tal14] A. Tal. *Shrinkage of De Morgan formulae by spectral techniques*. FOCS 2014, pages 551–560.

[Tam16] S. Tamaki. *A Satisfiability Algorithm for Depth Two Circuits with a Sub-Quadratic Number of Symmetric and Threshold Gates*. ECCC TR16-100, 2016.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il