Electronic Colloquium on Computational Complexity, Report No. 100 (2016)

# A Satisfiability Algorithm for Depth Two Circuits with a Sub-Quadratic Number of Symmetric and Threshold Gates

Suguru Tamaki*

**Abstract**

We consider depth 2 unbounded fan-in circuits with symmetric and linear threshold gates. We present a deterministic algorithm that, given such a circuit with $n$ variables and $m$ gates, counts the number of satisfying assignments in time $2^{n-\Omega\left(\left(\frac{n}{\sqrt{m}\cdot\mathrm{poly}(\log n)}\right)^{a}\right)}$ for some constant $a > 0$. Our algorithm runs in time super-polynomially faster than $2^n$ if $m = O(n^2/\log^b n)$ for some constant $b > 0$. Previously, such algorithms were only known for bounded depth circuits with linear threshold gates and a slightly super-linear number of *wires* [Impagliazzo-Paturi-Schneider, FOCS 2013 and Chen-Santhanam-Srinivasan, CCC 2016].

We also show that depth 2 circuits with $O(n^2/\log^b n)$ symmetric and linear threshold gates in total cannot compute an explicit function computable by a deterministic $2^{O(n)}$-time Turing machine with an NP oracle. Previously, even slightly super-linear lower bounds on the number of gates were not known until recently Kane and Williams [STOC 2016] showed that depth 2 linear threshold circuits with $o(n^{3/2}/\log^3 n)$ gates cannot compute an explicit function computable in linear time.

**Key words:** exponential time algorithm, circuit lower bound, polynomial method, derandomization

*Kyoto University `tamak@kuis.kyoto-u.ac.jp`. Supported in part by MEXT KAKENHI (24106003); JSPS KAKENHI (25240002, 26330011); the John Mung Advanced Program of Kyoto University. Part of the work performed while the author was at Department of Computer Science and Engineering, University of California, San Diego and the Simons Institute for the Theory of Computing, Berkeley.

ISSN 1433-8092

# 1 Introduction

We are concerned with circuits that consist of unbounded fan-in symmetric and linear threshold gates. Let $x_1, x_2, \dots, x_n$ be Boolean variables and $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. We say $f$ is *symmetric* if there exists a function $g : \mathbb{Z} \to \{0,1\}$ such that $f(x) = g(\sum_{i=1}^n x_i)$ holds. We say $f$ is a *linear threshold* function (LTF) if there exist $w_0, w_1, \dots, w_n \in \mathbb{Z}$ such that $f(x) = \mathrm{sgn}(w_0 + \sum_{i=1}^n w_i x_i)$ holds, where $\mathrm{sgn} : \mathbb{Z} \to \{0,1\}$ is the sign function defined as $\mathrm{sgn}(y) = 1$ if and only if $y \geq 0$.

In this paper, we present satisfiability algorithms and circuit size lower bounds for depth 2 circuits with symmetric and linear threshold gates as described in the next section. Note that each gate of such a circuit may be of a different type, e.g., $g_1(\sum_{i=1}^n x_i), g_2(\sum_{i=1}^n x_i), \dots, \mathrm{sgn}(w_{1,0} + \sum_{i=1}^n w_{1,i} x_i), \mathrm{sgn}(w_{2,0} + \sum_{i=1}^n w_{2,i} x_i), \dots$ etc.

## 1.1 Our contribution

**Satisfiability algorithms** In this paper, we present the following satisfiability algorithms.

**Theorem 1.1** (Main 1)**.** *There exist a constant $c > 0$ and a deterministic algorithm that, given a depth 2 circuit $C$ with $n$ variables and $m$ gates, where each gate is either symmetric or linear threshold, runs in time $2^{n-\Omega\left(\left(\frac{n}{\sqrt{m}\cdot \mathrm{poly}(\log n)}\right)^c\right)}$ and counts the number of satisfying assignments for $C$.*

Previously, Impagliazzo, Paturi and Schneider [33] showed that the satisfiability of a depth 2 linear threshold circuit with $n$ variables and $m$ *wires* can be solved in randomized time $2^{n-\mu(m/n)n}$, where $\mu(c) = 1/c^{O(c^2)}$. Chen and Santhanam [16] improved the running time as $\mu(c) = 1/c^{O(c)}$. Chen, Santhanam and Srinivasan [17] showed that the satisfiability of a depth $d$ linear threshold circuit with $n$ variables and $n^{1+\varepsilon_d}$ wires can be solved in randomized time $2^{n-n^{\varepsilon_d}}$, where $\varepsilon_d = 1/2^{O(d)}$.

Note that a depth 2 linear threshold circuit with $m$ gates may have $O(mn)$ wires. We are not aware of satisfiability algorithms that beat brute force search for depth 2 circuits with symmetric and linear threshold gates as Theorem 1.1 or even for depth 2 circuits with only symmetric gates. To summarize, our algorithm is deterministic, can solve a counting version of the satisfiability problem and handle larger size circuits (of depth 2) with additional gate types. Our algorithm can be generalized to handle bounded depth layered circuits, where each layer consists of either AND/OR/XOR gates or symmetric and linear threshold gates and the fan-in of symmetric and linear threshold gates satisfies some condition.

**Circuit lower bounds** As a byproduct of Theorem 1.1, we obtain the following circuit lower bounds.

**Theorem 1.2** (Main 2)**.** *There exist a language $L \in \mathrm{E}^{\mathrm{NP}}$ and a constant $c > 0$ such that any family of depth 2 circuits with $O(n^2/\log^c n)$ gates, where each gate is either symmetric or linear threshold, cannot compute $L$.*

Here $\mathrm{E}^{\mathrm{NP}}$ is the class of languages computable by deterministic $2^{O(n)}$-time NP-oracle Turing machines. It has been a longstanding open question whether $\mathrm{E}^{\mathrm{NP}}$ can be computed by depth 2 circuits with $n^{1.01}$ threshold gates until very recently Kane and Williams [35] showed that depth 2 circuits with $o(n^{3/2}/log^3 n)$ linear threshold gates cannot compute an explicit function computable in linear time.

Again we are not aware of non-trivial lower bounds for depth 2 circuits with symmetric and linear threshold gates as Theorem 1.2 or even for depth 2 circuits with only symmetric gates. To summarize, we show lower bounds for larger size circuits with additional gate types computing a less explicit function.

## 1.2 Background and Related Work

The motivation for studying the satisfiability problem of depth 2 linear threshold circuits is twofold: First, the problem contains as special cases both of the maximum satisfiability problem and 0-1 integer linear programming, which have been well studied in the area of exponential time algorithms and implementations of practical solvers. Second, proving super-polynomial lower bounds against depth 2 linear threshold circuits is one of the major open questions in Boolean circuit complexity. Below we elaborate on the second point.

Bounded depth linear threshold circuits have been studied extensively as a model of neural network. Such circuits are powerful enough to implement arithmetic operations such as iterated multiplication, division and powering, see, e.g., [41] and even candidate pseudorandom function generators [37, 42, 45]. The latter fact explains the difficulty of proving lower bounds for bounded depth linear threshold circuits by the "Natural Proof" barrier due to Razborov and Rudich [52, 18, 64] although it is believed that such circuits cannot compute some functions in NP or ever in P. There has been much effort to reveal the expressive power of linear threshold circuits, see, e.g., [26, 50, 32, 28, 27, 36, 39, 38, 20, 19, 22, 53, 47, 21, 25, 51, 40], to name a few.

The connection between satisfiability algorithms and circuit lower bounds, developed by Williams and subsequent authors [63, 67, 62, 64, 66, 9, 34], is a promising approach to avoid such barriers, see also [56, 49, 65] for surveys. Since the success of using the connection to actually prove new circuit lower bounds, i.e., super-polynomial lower bounds for $\mathbf{ACC}^0$ circuits computing a language in NEXP [67], many satisfiability algorithms that beat brute force search have been designed for various circuit classes [48, 66, 15, 3, 2, 44, 59]. Interestingly, some papers showed average-case circuit lower bounds directly from the analyses of their satisfiability algorithms [55, 31, 4, 57, 14, 13, 12, 23, 54].

## 1.3 Techniques

The polynomial method is a powerful technique in Boolean circuit complexity [5]. In his remarkable result, Williams [67] used the polynomial method to design satisfiability algorithms beating brute force search for $\mathbf{ACC}^0$ circuits. Since then, Williams and his coauthors have developed algorithms for many interesting problems such as the circuit satisfiability problem for restricted classes of circuits [66], all-pairs shortest paths [11] and Hamming nearest neighbors [1], see also [68].

We follow the approach of [66] that gives satisfiability algorithms beating brute force search for $\mathbf{ACC}^0 \circ \mathbf{THR}$ and $\mathbf{ACC}^0 \circ \mathbf{SYM}$ circuits. The approach is summarized as follows: (1) Given an $n$-variate circuit $C \in \mathbf{ACC}^0 \circ \mathbf{THR}$, consider a circuit $C'(y) := \bigvee_{a \in \{0,1\}^{n'}} C(y,a)$ for some $n' < n$. (2) Represent $C'$ as a circuit in $\mathbf{SYM} \circ \mathbf{SYM}$ using simulation techniques, in particular, the simulation of Beigel and Tarui [8] that transforms a circuit in $\mathbf{ACC}^0$ to a circuit in $\mathbf{SYM} \circ \mathbf{AND}$. (3) Apply the "fast evaluation algorithm" for $\mathbf{SYM} \circ \mathbf{SYM}$ to obtain the truth table of $C'$.

We implement the above approach, focusing on Item (2), for $\mathbf{THR} \circ \mathbf{THR}$ circuits. If we use the construction of "probabilistic polynomials" for symmetric and linear threshold functions due to Srinivasan [58], we can represent $C'$ as a "probabilistic circuit" in $\mathbf{SYM} \circ \mathbf{SYM}$. This implementation of Item (2) is sufficient to obtain randomized algorithms.

In order to design deterministic algorithms, we derandomize probabilistic polynomials of [58]. It turns out that *pseudorandom generators for space-bounded computation* due to Nisan [46] is sufficient for our purpose. We also use *modulus-amplifying polynomials* [61, 69] to complete a deterministic implementation of Item (2) in a similar way to [8, 11].

Our circuit lower bounds follow from the connection between satisfiability algorithms and circuit lower bounds, in particular, the one due to Ben-Sasson and Viola [9].

# 2 Preliminaries

We use the following notations: $\mathbb{Z}$ is the set of integers, $\mathbb{N}$ is the set of natural numbers, i.e., non-negative integers, $\mathbb{Z}_m$ is the quotient ring of integers modulo $m$, identified with $\{0,1,\dots,m-1\}$, $\mathbb{F}_2$ is the finite field of order 2, identified with $\{0,1\}$.

For a positive integer $n$, $[n] := \{1,2,\dots,n\}$. For real numbers $a<b$, $(a,b)$ is the open interval between $a$ and $b$. For $y \in \mathbb{Z}$, $|y|$ is the absolute value of $y$. For a finite set $S$, $|S|$ is the cardinality of $S$. For $x \in \{0,1\}^n$, $|x|$ is the *Hamming weight* of $x$, i.e., $|x| = \sum_{i=1}^n x_i$.

The logarithm of $x$ to base 2 is $\lg x$ and that to base $e$ is $\ln x$. We use random access machines as our computation model.

## 2.1 Probability and derandomization

We use the following results in Section 3.1.

**Lemma 2.1** (The Chernoff-Hoeffding bound [29]). *Let $X_1,\dots,X_n$ be independent and identically distributeid Bernoulli random variables with $\mathbf{Pr}[X_i=1]=1-\mathbf{Pr}[X_i=0]=1/m$. Then, it holds that*

$$\mathbf{Pr}[|n/m-\sum_{i=1}^n X_i|>t] \le 2e^{-2t^2/n}.$$

**Lemma 2.2** (Nisan [46]). *Let $f:\mathbb{Z}_m^n \to \{0,1\}$ be a function computable in space $O(\lg(n\lg m))$. Then, there exists a function $G:\{0,1\}^\ell \to \mathbb{Z}_m^n$ with $\ell = O(\lg^2((n\lg m)/\varepsilon))$ such that*

- *$|\mathbf{Pr}[f(x)=1]-\mathbf{Pr}[f(G(y))=1]| \le \varepsilon$, where $x$ and $y$ are respectively sampled from $\mathbb{Z}_m^n$ and $\{0,1\}^\ell$ uniformly at random, and*
- *$G$ is computable in time $\mathrm{poly}(n\lg m)$.*

## 2.2 Boolean circuits

Let $x_1,x_2,\dots,x_n$ be Boolean variables and $f:\{0,1\}^n \to \{0,1\}$ be a Boolean function. We say $f$ is $W$-*sum* if there exist a function $g:\mathbb{Z}\to\{0,1\}$ and $w_1,\dots,w_n \in \mathbb{N}$ with $\sum_{i=1}^n w_i \le W$ such that $f(x)=g(\sum_{i=1}^n w_i x_i)$ holds. Note that we can realize a $W$-sum function as a $W$-variate symmetric function by regarding $w_i x_i$ as a sum of $w_i$ variables. In what follows, we identify Boolean functions and logic gates.

We denote by $\mathbf{AND},\mathbf{OR},\mathbf{XOR},\mathbf{SYM},\mathbf{SUM}_W,\mathbf{THR}$ the set of AND gates, the set of OR gates, the set of XOR gates, the set of symmetric gates, the set of $W$-sum gates, the set of linear threshold gates, respectively. Let $\mathscr{G}_0,\mathscr{G}_1,\dots,\mathscr{G}_{d-1} \in \{\mathbf{AND},\mathbf{OR},\mathbf{XOR},\mathbf{SYM},\mathbf{SUM}_W,\mathbf{THR}\}$ be sets of logic gates. We denote by $\mathscr{G}_0 \circ \mathscr{G}_1 \circ \cdots \circ \mathscr{G}_{d-1}$ the set of depth $d$ unbounded-fan-in layered Boolean circuits such that layer $i$ contains gates from $\mathscr{G}_i$ and all the gates at layer $i$ are only fed by gates at layer $i+1$. Layer 0 corresponds to the output gate and layer $d$ consists of input variables and constants $0,1$. We allow inputs and outputs of gates to be negated unless otherwise specified.

We need the following upper bounds on the weights of linear threshold functions in Section 3.3.

**Lemma 2.3** (Muroga [43]). *For all $w_0,w_1,\dots,w_n \in \mathbb{Z}$, there exist $w_0',w_1',\dots,w_n' \in \mathbb{Z}$ with $|w_i'| = 2^{O(n\lg n)}$ such that $\mathrm{sgn}(w_0+\sum_{i=1}^n w_i x_i) = \mathrm{sgn}(w_0'+\sum_{i=1}^n w_i' x_i)$ holds. In addition, $w_0',w_1',\dots,w_n'$ can be efficiently obtained.*

We use the following results in Section 4.

**Lemma 2.4** (Maciel-Thérien [41], see also Section 2.2 in the arXiv version of Williams [66])**.** *There exists a positive integer $c_{\mathbf{mt}}$ such that for all $n$-variate $f \in \mathbf{THR}$, there exists a circuit*

$$C \in \mathbf{OR} \circ \mathbf{AND} \circ \mathbf{XOR} \circ \mathbf{OR} \circ \mathbf{AND} \circ \mathbf{SYM}$$

*that is equivalent to $f$ and consists of at most $n^{c_{\mathbf{mt}}}$ wires.*

**Lemma 2.5** (Beigel [6])**.** *For all circuit $C \in \mathbf{AND} \circ \mathbf{SYM}$ whose AND gate at layer $0$ has fan-in $t_1$ and symmetric gates at layer $1$ have fan-in at most $t_2$, there exists a circuit $C' \in \mathbf{SYM}$ that is equivalent to $C$ and whose fan-in is at most $(t_2+1)^{t_1}$.*

**Lemma 2.6** (Williams [66])**.** *There exists a positive constant $c_{\mathbf{w}}$ and an algorithm that, given an $n$-variate circuit $C \in \mathbf{SYM} \circ \mathbf{SYM}$ whose symmetric gate at layer $0$ has fan-in at most $t_1$ and symmetric gates at layer $1$ have fan-in at most $t_2$ such that $t_1 t_2 \leq 2^{c_{\mathbf{w}} n}$, prints the truth table of $C$ in time $\mathrm{poly}(n)2^n$.*

## 2.3 Polynomials

Let $x_1, x_2, \ldots, x_n$ be formal variables and $K \in \{\mathbb{F}_2, \mathbb{Z}\}$. In this paper, each variable always takes the values 0 or 1, hence the identity $x_i^2 = x_i$ holds. A *monomial* is a product of variables, i.e., $\prod_{i \in S} x_i$ for some $S \subseteq [n]$. For $S = \emptyset$, we regard $\prod_{i \in S} x_i$ as 1. We can represent a *$K$-polynomial* $P$ as a sum of terms, of the form $P(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i$, where $a_S \in K$. Whenever we consider a $\mathbb{Z}$-polynomial, $|a_S| = 2^{O(n)}$ is assumed unless otherwise stated. The *degree* of $P$, denoted by $\deg(P)$, is defined as

$$\deg(P) := \max\{|S| \mid S \subseteq [n], a_S \neq 0\}.$$

Note that we can regard an $\mathbb{F}_2$-polynomial $P$ as a Boolean circuit in $\mathbf{XOR} \circ \mathbf{AND}$ as

$$P(x) = \bigoplus_{S \subseteq [n]} \left( a_S \bigwedge_{i \in S} x_i \right).$$

We need the following combinatorial facts in Sections 3 and  4.

**Lemma 2.7.** *The number of monomials of degree at most $k$ is $M(n,k) = \sum_{i=0}^{k} \binom{n}{i}$. If $k \leq n/2$, $M(n,k) \leq k\binom{n}{k}$.*

**Lemma 2.8** (Powering)**.** *Given an $n$-variate degree $k$ polynomial $P$ represented as a sum of terms, and a positive integer $d$, we can represent $P^d$, the $d$th power of $P$, as a sum of terms in time*

$$\mathrm{poly}(n) \sum_{i=1}^{d-1} M(n,k)M(n,ik) \leq \mathrm{poly}(n)M(n,dk).$$

**Lemma 2.9** (Composition)**.** *Let $p$ be a degree $d_1$ polynomial in $n_1$ variables and $p_1, p_2, \ldots, p_{n_1}$ be degree $d_2$ polynomials in the same $n_2$ variables. Then, $p(p_1, p_2, \ldots, p_{n_1})$ can be represented as a sum of terms in time* $\mathrm{poly}(n_1, n_2)M(n_1, d_1)M(n_2, d_1 d_2)$.

We use the following construction of polynomials approximating symmetric functions in Section 3.1.

**Lemma 2.10** (Corollary 2.7 in Bhatnagar-Gopalan-Lipton [10], Lemma 3.1 in Alman-Williams [1])**.** *For all $n$-variate function $f \in \mathbf{SYM}$ and integers $s \geq 0, t \geq 1$ with $s + t \leq n$, there exists an $\mathbb{F}_2$-polynomial $p$ of degree at most $O(t)$ such that $f(x) = p(x)$ holds if $s \leq |x| \leq s+t$. In addition, $p$ can be constructed in time* $\mathrm{poly}(n)\binom{n}{O(t)}$.

We need the following construction of *modulus-amplifying polynomials* [61, 69] in Section 4.

**Lemma 2.11** (Beigel-Tarui [8]). *For every positive integer $\ell$, the degree $(2\ell-1)$ univariate $\mathbb{Z}$-polynomial*

$$F_\ell(y) := 1-(1-y)^\ell \sum_{j=0}^{\ell-1} \binom{\ell+j-1}{j} y^j$$

*satisfies:*

- *if $y = 0 \bmod 2$, then $F_\ell(y) = 0 \bmod 2^\ell$,*
- *if $y = 1 \bmod 2$, then $F_\ell(y) = 1 \bmod 2^\ell$.*

*In addition, for $0 \le i \le 2\ell-1$, the coefficient of $y^i$ in the polynomial $F_\ell$ has magnitude at most $2^{O(\ell)}$.*

### 2.4 Probabilistic polynomials

For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, a probability distribution $\mathscr{P}$ over polynomials is an *$\varepsilon$-error probabilistic polynomial* for $f$ if for all $x \in \{0,1\}^n$, $\mathbf{Pr}_{p\sim\mathscr{P}}[f(x) \neq p(x)] \le \varepsilon$ holds [60]. The *degree* of a probabilistic polynomial $\mathscr{P}$ is the maximum degree of polynomials in the support of $\mathscr{P}$, i.e., $\max\{\deg(p) \mid \mathbf{Pr}_{q\sim\mathscr{P}}[p=q] > 0\}$. A probabilistic polynomial $\mathscr{P}$ has *$r$-randomness* if we can sample a polynomial from $\mathscr{P}$ with $r$ uniformly random bits.

We need the following construction of time and randomness efficient probabilistic polynomials for AND/OR functions in Section 4.

**Lemma 2.12** (Beigel-Reingold-Spielman and Tarui [7, 60]). *For every $\varepsilon \in (0,1/2)$, there exists an $\varepsilon$-error probabilistic $O(\lg^2 n \cdot \lg(1/\varepsilon))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}$ of degree $d = O(\lg n \cdot \lg(1/\varepsilon))$ for $n$-variate AND/OR functions. Furthermore, we can sample a polynomial from $\mathscr{P}$ in time $O(\mathrm{poly}(n)\binom{n}{d})$.*

## 3 Randomness efficient probabilistic polynomials

In this section, we present the main technical ingredients of our satisfiability algorithms, that is, a time and randomness efficient version of probabilistic polynomials for weighted symmetric and linear threshold functions due to Srinivasan [58].

**Lemma 3.1** (Randomness efficient version of Theorem 11 in [58]). *For every $\varepsilon \in (0,1/2)$, $W \in \mathbb{N}$ and an $n$-variate $f \in \mathbf{SUM}_W$, $f$ has an $\varepsilon$-error $O(\lg^2((n\lg\lg W)/\varepsilon))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}$ of degree $d = O(\lg^4 W\sqrt{n\lg(1/\varepsilon)})$. Furthermore, we can sample a polynomial from $\mathscr{P}$ in time $O(\mathrm{poly}(n)\binom{n}{d})$.*

**Lemma 3.2** (Randomness efficient version of Theorem 12 in [58]). *For every $\varepsilon \in (0,1/2)$ and an $n$-variate $f \in \mathbf{THR}$, $f$ has an $\varepsilon$-error $O(\lg^2(n/\varepsilon))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}$ of degree $d = O(\lg^5 n\sqrt{n\lg(1/\varepsilon)})$. Furthermore, we can sample a polynomial from $\mathscr{P}$ in time $O(\mathrm{poly}(n)\binom{n}{d})$.*

Lemma 3.3 below is the key result of this section. First we need some definitions. For $m \in \mathbb{N}, r \in \mathbb{Z}_m, w \in \mathbb{Z}_m^n$, we define functions $\mathrm{mod}_{m,r}^n : \{0,1\}^n \to \{0,1\}$, $\mathrm{mod}_{m,r,w}^n : \{0,1\}^n \to \{0,1\}$, as follows.

- $\mathrm{mod}_{m,r}^n(x) = 1$ if and only if $\sum_{i=1}^n x_i \equiv r \bmod m$,
- $\mathrm{mod}_{m,r,w}^n(x) = 1$ if and only if $\sum_{i=1}^n w_i x_i \equiv r \bmod m$.

**Lemma 3.3** (Randomness efficient version of Lemma 13 in [58]). *For every $\varepsilon \in (0,1/2)$, $\mathrm{mod}_{m,r,w}^n$ has an $\varepsilon$-error $O(\lg^2((n\lg m)/\varepsilon))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}$ of degree $d = O(m\sqrt{n\lg(1/\varepsilon)})$. Furthermore, we can sample a polynomial from $\mathscr{P}$ in time $O(\mathrm{poly}(n)\binom{n}{d})$.*

We prove the above lemma in the next section. The proof is based on the observation that uniformly random bits in the construction of [58] can be replaced by the outputs of the pseudorandom generators for space-bounded computation due to Lemma 2.2.

Once we establish Lemma 3.3, we can prove Lemmas 3.1 and 3.2 following the lead of [58] with careful calculation of parameters. The proofs are given in Sections 3.2 and 3.3 respectively.

## 3.1 Weighted modulo functions

In this section, we prove Lemma 3.3.

Fix integers $m \geq 2$ and $r \in \mathbb{Z}_m$ and an integer vector $w \in \mathbb{Z}_m^n$. Let $v \in \mathbb{Z}_m^n$. We define functions $M_{m,r} : \mathbb{Z}_m \to \{0,1\}$, $M_{m,r,w,v}^n : \{0,1\}^n \to \{0,1\}^n$ and a set $R_{m,r,v}^n \subseteq \mathbb{Z}_m^{m-1}$ as follows.

- $M_{m,r}(y) = 1$ if and only if $y \equiv r \bmod m$,
- $(M_{m,r,w,v}^n(x))_i := M_{m,r}(w_i x_i + v_i)$,
- $R_{m,r,v}^n := \{(r_1, r_2, \ldots, r_{m-1}) \in \mathbb{Z}_m^{m-1} \mid \sum_{i=1}^{m-1} ir_i \equiv r + \sum_{i=1}^n v_i \bmod m\}$.

Note that $M_{m,r}(w_i x_i + v_i) \in \{0, 1, x_i, 1 - x_i\}$ holds for fixed $m, r, w_i, v_i$. The following lemma shows how to reduce the evaluation of $\mathrm{mod}_{m,r,w}^n(x)$ to the evaluation of $\mathrm{mod}_{m,r'}^n(x')$ for many pairs $(r', x')$.

**Lemma 3.4** (Section 3.1 in [58])**.** *For all $v \in \mathbb{Z}_m^n$ and $x \in \{0,1\}^n$, it holds that*

$$\mathrm{mod}_{m,r,w}^n(x) = \sum_{u \in R_{m,r,v}} \bigwedge_{i=1}^{m-1} \mathrm{mod}_{m,u_i}^n(M_{m,r,w,v}^n(x)).$$

Let $P_{m,r}^n : \{0,1\}^n \to \{0,1\}$ be an $\mathbb{F}_2$-polynomial of degree $O(t)$ such that $P_{m,r}^n(x) = \mathrm{mod}_{m,r}^n(x)$ if $|x| \in \{\lfloor n/m \rfloor - t, \ldots, \lfloor n/m \rfloor + t\}$. By Lemma 2.10, the existence of $P_{m,r}^n$ is guaranteed. In addition, $P_{m,r}^n$ can be constructed in time $\mathrm{poly}(n)\binom{n}{O(t)}$. Let us define an $\mathbb{F}_2$-polynomial $Q_{m,r,w,v}^n : \{0,1\}^n \to \{0,1\}$ as follows.

$$Q_{m,r,w,v}^n(x) := \sum_{u \in R_{m,r,v}} \prod_{i=1}^{m-1} P_{m,u_i}^n(M_{m,r,w,v}^n(x)).$$

The following lemma is immediate from the property of $P_{m,r}^n$ and the definition of $Q_{m,r,w,v}^n$.

**Lemma 3.5.** *If $|M_{m,r,w,v}^n(x)| \in \{\lfloor n/m \rfloor - t, \ldots, \lfloor n/m \rfloor + t\}$, then $Q_{m,r,w,v}^n(x) = \mathrm{mod}_{m,r,w}^n(x)$ holds.*

We are ready to prove Lemma 3.3.

*Proof of Lemma 3.3.* If we select $v_i \in \mathbb{Z}_m$ uniformly at random, then we have $\mathbf{Pr}_{v_i}[M_{m,r}(w_i x_i + v_i) = 1] = 1/m$. Hence, if we select $v \in \mathbb{Z}_m^n$ uniformly at random, then by Lemma 2.1, we have

$$\mathbf{Pr}_{v}[|M_{m,r,w,v}^n(x)| \notin \{\lfloor n/m \rfloor - t, \ldots, \lfloor n/m \rfloor + t\}] \leq 2e^{-2t^2/n}.$$

Let $\ell = O(\lg^2((n \lg m)/\delta))$ and $G : \{0,1\}^\ell \to \mathbb{Z}_m^n$ be the pseudorandom generator due to Lemma 2.2. Since $|M_{m,r,w,v}^n(x)|$ as a function of $v$ can be computed in space $O(n \lg m)$, if we select $s \in \{0,1\}^\ell$ uniformly at random, then we have

$$\mathbf{Pr}_{s}[|M_{m,r,w,G(s)}^n(x)| \notin \{\lfloor n/m \rfloor - t, \ldots, \lfloor n/m \rfloor + t\}] \leq 2e^{-2t^2/n} + \delta.$$

This implies

$$\mathbf{Pr}_{s}[Q_{m,r,w,G(s)}^n(x) \neq \mathrm{mod}_{m,r,w}^n(x)] \leq 2e^{-2t^2/n} + \delta.$$

If we set $t = \sqrt{(n/2)\ln(4/\varepsilon)}$ and $\delta = \varepsilon/2$, then the right hand side is at most $\varepsilon$ and the degree of $Q_{m,r,w,G(s)}^n(x)$ is $O(tm)$. This completes the proof. □

### 3.2 Weighted sum functions

In this section, we prove Lemma 3.1.

Fix a function $g : \mathbb{Z} \to \{0,1\}$ and natural numbers $w_1, \dots, w_n$ with $\sum_{i=1}^n w_i = W$. Let $f(x) = g(\sum_{i=1}^n w_i x_i)$, $\ell := \lceil \lg W \rceil + 2$, $p_1 < \cdots < p_\ell$ be first $\ell$ primes and $s := \sum_{i=1}^{\ell} p_i$.

Note that $\prod_{i=1}^{\ell} p_i > 2^\ell > 2W$. By the prime number theorem, $p_\ell = O(\lg W \cdot \lg\lg W)$ holds and this implies $s = O(\lg^2 W \cdot \lg\lg W)$.

We define functions $M^n_{m,w} : \{0,1\}^n \to \{0,1\}^m$ for $m \in \mathbb{N}$ and $M^n_w : \{0,1\}^n \to \{0,1\}^s$ as follows.

- $M^n_{m,w}(x) := (\mathrm{mod}^n_{m,0,w}(x), \dots, \mathrm{mod}^n_{m,m-1,w}(x))$,
- $M^n_w(x) := (M^n_{p_1,w}(x), \dots, M^n_{p_\ell,w}(x))$.

Since we can reconstruct $\sum_{i=1}^n w_i x_i$ from $M^n_w(x)$ by the Chinese remainder theorem, we have:

**Lemma 3.6** (Section 3.2 in [58]). *There exists a function $h : \{0,1\}^s \to \{0,1\}$ such that $f(x) = h(M^n_w(x))$ holds.*

Note that $h$ can be written as an $\mathbb{F}_2$-polynomial of degree at most $s$ and is determined by the values $g(0), g(1), \dots, g(W)$. We are ready to prove Lemma 3.1.

*Proof of Lemma 3.1.* For each $p_i$ and $r \in \mathbb{Z}_{p_i}$, there exists a $\delta$-error $O(\lg^2((n \lg p_i)/\delta))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}_{p_i,r}$ of degree $O(p_i \sqrt{n \lg(1/\delta)})$ by Lemma 3.3. We sample an $\mathbb{F}_2$-polynomial $P_{p_i,r}$ from $\mathscr{P}_{p_i,r}$, replace $\mathrm{mod}^n_{p_i,r,w}$ by it in $M^n_w$ and then obtain a polynomial $Q$ for $f$ by composing $h$. Note that we use same random bits of length at most $O(\lg^2((n \lg p_\ell)/\delta))$ to sample every $P_{p_i,r}$.

By the union bound, we have $\mathbf{Pr}[Q(x) \neq f(x)] \leq s\delta$. If we set $\delta = \varepsilon/s$, the degree of $Q$ is $O(s p_\ell \sqrt{n \lg(1/\delta)}) = O(\lg^4 W \sqrt{n \lg(1/\varepsilon)})$ and the length of random bits is $O(\lg^2((n \lg\lg W)/\varepsilon))$. This completes the proof. □

### 3.3 Linear threshold functions

In this section, we prove Lemma 3.2.

Fix integers $w_0, w_1, \dots, w_n \in \mathbb{N}$, let $F(x) = w_0 + \sum_{i=1}^n w_i x_i$ and consider $\mathrm{sgn}(F(x)) \in \mathbf{THR}$. Without loss of generality, $|w_i| \leq 2^{O(n \lg n)}$ holds due to Lemma 2.3. We assume that $|F(x)| \geq n+2$. Otherwise, we consider $(n+2)(2F(x)+1)$ instead since for all $x \in \{0,1\}^n$, it holds that $\mathrm{sgn}(F(x)) = \mathrm{sgn}((n+2)(2F(x)+1))$ and $|(n+2)(2F(x)+1)| \geq n+2$.

Let $\ell := \lceil \lg((n+1)\max_i |w_i|) \rceil$. We need the following definitions for $1 \leq l \leq \ell$.

- $w_i^{(l)} := \begin{cases} \lfloor w_i/2^l \rfloor & \text{if } w_i \geq 0, \\ w_i^{(l)} = \lceil w_i/2^l \rceil & \text{if } w_i < 0, \end{cases}$
- $F^{(l)}(x) := w_0^{(l)} + \sum_{i=1}^n w_i^{(l)} x_i$,
- $\mathrm{ins}^{(l)}(x) = 1$ if and only if $w_0^{(l)} + \sum_{i=1}^n w_i^{(l)} x_i \in \{-n-1, -n, \dots, n, n+1\}$,
- $\mathrm{pos}^{(l)}(x) = 1$ if and only if $w_0^{(l)} + \sum_{i=1}^n w_i^{(l)} x_i \in \{0, 1, \dots, n, n+1\}$,
- $\mathrm{ins}_p^{(l)}(x) = 1$ if and only if $w_0^{(l)} + \sum_{i=1}^n w_i^{(l)} x_i \equiv k \bmod p$ for some $k \in \{-n-1, -n, \dots, n, n+1\}$,
- $\mathrm{pos}_p^{(l)}(x) = 1$ if and only if $w_0^{(l)} + \sum_{i=1}^n w_i^{(l)} x_i \equiv k \bmod p$ for some $k \in \{0, 1, \dots, n, n+1\}$,

Hofmeister gives the following characterization of linear threshold functions.

**Lemma 3.7** (page 139, [30])**.** *If $F(x) \geq 0$, then there exists a unique $l$ such that $\neg \mathrm{ins}^{(l-1)}(x) \wedge \mathrm{pos}^{(l)}(x) = 1$ holds. If $F(x) < 0$, then for all $l$, $\neg \mathrm{ins}^{(l-1)}(x) \wedge \mathrm{pos}^{(l)}(x) = 0$ holds.*

The following lemma implies Lemma 3.2 almost immediately.

**Lemma 3.8.** *For every $\varepsilon \in (0, 1/2)$ and $l$, $f \in \{\mathrm{ins}^{(l)}, \mathrm{pos}^{(l)}\}$ has an $\varepsilon$-error $O(\lg^2(n/\varepsilon))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}$ of degree $d = O(\lg^4 n\sqrt{n \lg(1/\varepsilon)})$. Furthermore, we can sample a polynomial from $\mathscr{P}$ in time $O(\mathrm{poly}(n)\binom{n}{d})$.*

First we prove Lemma 3.2 assuming Lemma 3.8 and then prove Lemma 3.8.

*Proof of Lemma 3.2.* For each $l$, there exist $\delta$-error $O(\lg^2(n/\delta))$-randomness probabilistic $\mathbb{F}_2$-polynomials $\mathscr{P}_{\mathrm{ins}}^{(l)}$ and $\mathscr{P}_{\mathrm{pos}}^{(l)}$ of degree $O(\lg^4 n\sqrt{n \lg(1/\delta)})$ for $\mathrm{ins}^{(l)}$ and $\mathrm{pos}^{(l)}$ respectively by Lemma 3.8. We sample an $\mathbb{F}_2$-polynomial $P_{\mathrm{ins}}^{(l)}$ from $\mathscr{P}_{\mathrm{ins}}^{(l)}$ and an $\mathbb{F}_2$-polynomial $P_{\mathrm{pos}}^{(l)}$ from $\mathscr{P}_{\mathrm{pos}}^{(l)}$ and construct an $\mathbb{F}_2$-polynomial $P(x) := \sum_{l=1}^{\ell}(1 - P_{\mathrm{ins}}^{(l-1)(x)})P_{\mathrm{pos}}^{(l)}(x)$. Note that we use same random bits of length at most $O(\lg^2(n/\delta))$ to sample every $P_{\mathrm{ins}}^{(l)}, P_{\mathrm{pos}}^{(l)}$.

By the union bound, we have $\mathbf{Pr}[P(x) \neq f(x)] \leq 2\ell\delta$. If we set $\varepsilon = 2\ell\delta$, the degree of $Q$ is $O(\lg^5 n\sqrt{n \lg(1/\varepsilon)})$ and the length of random bits is $O(\lg^2(n/\varepsilon))$.

This completes the proof of Lemma 3.2. □

*Proof of Lemma 3.8.* We show a proof for $\mathrm{ins}^{(l)}$. The proof for $\mathrm{pos}^{(l)}$ is almost identical. The main idea is that we compute $\mathrm{ins}_p^{(l)}$ instead of $\mathrm{ins}^{(l)}$ for a random prime $p$. Note that $\mathrm{ins}_p^{(l)} \in \mathbf{SUM}_W$ for $W \leq pn$. There exists an $\varepsilon$-error $O(\lg^2((n \lg\lg W)/\delta))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}_p^{(l)}$ of degree $O(\lg^4 W\sqrt{n \lg(1/\delta)})$ for $\mathrm{ins}_p^{(l)}$ by Lemma 3.1.

Let $t := \lceil Cn^2 \lg n/\delta \rceil$ for a sufficiently large constant $C > 0$ and $p_1 < \cdots < p_t$ be first $t$ primes. Note that $p_t = O(t \lg t)$ by the prime number theorem. We rely on the following lemma.

**Lemma 3.9** (Section 3.3 in [58])**.** *If $\mathrm{ins}^{(l)}(x) = 1$, then $\mathrm{ins}_p^{(l)}(x) = 1$. If $\mathrm{ins}^{(l)}(x) = 0$ and $i$ is selected from $\{1, 2, \ldots, t\}$ uniformly at random, then $\mathbf{Pr}_i[\mathrm{ins}_{p_i}^{(l)}(x) = 1] \leq \delta$.*

We construct an $\mathbb{F}_2$-polynomial $Q$ for $\mathrm{ins}^{(l)}$ as follows. First, select $i \in \{1, 2, \ldots, t\}$ uniformly at random. Then, sample a polynomial $P$ from $\mathscr{P}_{p_i}^{(l)}$ and let $Q(x) := P(x)$.

By the union bound, we have

$$\mathbf{Pr}[Q(x) \neq \mathrm{ins}^{(l)}(x)] \leq \mathbf{Pr}[\mathrm{ins}_{p_i}^{(l)}(x) \neq \mathrm{ins}^{(l)}(x)] + \mathbf{Pr}[P(x) \neq \mathrm{ins}_{p_i}^{(l)}(x)] \leq 2\delta.$$

If we set $\varepsilon = 2\delta$, the degree of $Q$ is $O(\lg^4 n\sqrt{n \lg(1/\varepsilon)})$ and the length of random bits is $O(\lg^2(n/\varepsilon))$.

This completes the proof of Lemma 3.8. □

# 4 Satisfiability Algorithms

In this section, we prove the following theorem.

**Theorem 4.1.** *There exist a constant $c > 0$ and a deterministic algorithm that, given a depth 2 linear threshold circuit $C$ with $n$ variables and $m$ gates, runs in time $2^{n-\Omega\left(\left(\frac{n}{\sqrt{m}\cdot\mathrm{poly}(\lg n)}\right)^c\right)}$ and counts the number of satisfying assignments for $C$.*

**Remark 4.2.** *The proof of Theorem 1.1 is essentially the same or even simpler and omitted, i.e., (1) we use Lemma 3.1 instead of 3.2 if necessary and (2) we do not have to apply Lemma 2.4 if a gate at the bottom layer is symmetric.*

Let $C \in \mathbf{THR} \circ \mathbf{THR}$ be an $n$-variate circuit whose gate at layer 0 has fan-in at most $m$. For a positive integer $n'$, we define a function $K : \{0,1\}^{n-n'} \to \{0,1,\ldots,2^{n'}\}$ as $K(y) := \sum_{a \in \{0,1\}^{n'}} C(y,a)$. Our goal is to construct an expression $K' = \sum_i a_i G_i$, where $a_i \in \mathbb{Z}$, $G_i \in \mathbf{SYM}$, such that $K \equiv K'$. Then $i$th bit of the binary representation of $K'(y) \in \{0,1\}^{n'+1}$ can be regarded as a function in $\mathbf{SYM} \circ \mathbf{SYM}$. We can apply Lemma 2.6 to obtain all the values of $K(y)$ if we select the underlying parameters appropriately.

*Proof of Theorem 4.1.* By Lemma 2.4, there exists $C' \in \mathbf{THR} \circ \mathbf{OR} \circ \mathbf{AND} \circ \mathbf{XOR} \circ \mathbf{OR} \circ \mathbf{AND} \circ \mathbf{SYM}$ that is equivalent to $C$ and has at most $t = mn^{c\mathbf{mt}}$ wires. Let $g_1, g_2, \ldots, g_s$ be symmetric gates at the bottom layer in $C'$. Let $C'' \in \mathbf{THR} \circ \mathbf{OR} \circ \mathbf{AND} \circ \mathbf{XOR} \circ \mathbf{OR} \circ \mathbf{AND}$ be an $s$-variate circuit with at most $t$ wires such that $C''(g_1, \ldots, g_s) \equiv C'$.

**Lemma 4.3.** *Let $D \in \mathbf{THR} \circ \mathbf{OR} \circ \mathbf{AND} \circ \mathbf{XOR} \circ \mathbf{OR} \circ \mathbf{AND}$ be an $n$-variate circuit with $t = \mathrm{poly}(n)$ wires, where the threshold gate at layer $0$ has fan-in at most $m = O(n^2)$. There exists an $\varepsilon$-error $O(\lg n \lg^2(n/\varepsilon))$-randomness probabilistic $\mathbb{F}_2$-polynomial $\mathscr{P}$ of degree $d = O(\lg^9 n \lg^5(1/\varepsilon)\sqrt{m})$ for $D$. Furthermore, we can sample from $\mathscr{P}$ in time $O(\mathrm{poly}(n)\binom{n}{d})$.*

*Proof.* We replace the threshold gate at layer 0 by a $\delta$-error probabilistic $\mathbb{F}_2$-polynomial from Lemma 3.2 and replace each AND/OR gate by a $\delta$-error probabilistic $\mathbb{F}_2$-polynomial from Lemma 2.12, where we set $\delta = \varepsilon/(t+1)$, and obtain a circuit

$$D' \in (\mathbf{XOR} \circ \mathbf{AND}) \circ (\mathbf{XOR} \circ \mathbf{AND}) \circ (\mathbf{XOR} \circ \mathbf{AND}) \circ \mathbf{XOR} \circ (\mathbf{XOR} \circ \mathbf{AND}) \circ (\mathbf{XOR} \circ \mathbf{AND}).$$

Note that we use the same random bits to sample each probabilistic polynomial. By repeatedly using Lemma 2.9, we obtain a circuit $D'' \in \mathbf{XOR} \circ \mathbf{AND}$ that is equivalent to $D'$. By the union bound, $D'$ is an $\varepsilon$-error probabilistic $\mathbb{F}_2$-polynomial for $D$. The degree of $D''$ is $d = O(\lg^9 n \lg^5(1/\varepsilon)\sqrt{m})$ and the randomness of $D'$ is $O(\lg n \lg^2(n/\varepsilon))$ by the choice of $\delta$. In addition, the construction of $D''$ takes time $O(\mathrm{poly}(n)\binom{n}{d})$ since we apply Lemma 2.9 at most

1. $t$ times with $d_1 = d_2 = O(\lg n \lg(1/\varepsilon)), n_1 = t, n_2 = n,$
2. $t$ times with $d_1 = O(\lg n \lg(1/\varepsilon)), d_2 = O(\lg^2 n \lg^2(1/\varepsilon)), n_1 = t, n_2 = n,$
3. $m$ times with $d_1 = O(\lg n \lg(1/\varepsilon)), d_2 = O(\lg^3 n \lg^3(1/\varepsilon)), n_1 = t, n_2 = n,$
4. once with with $d_1 = O(\lg^5 n \sqrt{m \lg(1/\varepsilon)}), d_2 = O(\lg^4 n \lg^4(1/\varepsilon)), n_1 = m, n_2 = n.$

This completes the proof. □

Let $l = O(\lg n \lg^2(n/\varepsilon)$ and select $r \in \{0,1\}^l$ to sample a polynomial $P_r$ for $C''$ due to Lemma 4.3 in time $\mathrm{poly}(n)\binom{s}{d_1}$, where $d_1 = O(\lg^9 n \lg^5(1/\varepsilon)\sqrt{m})$. Then we construct a $\mathbb{Z}$-polynomial $Q_r := F_\ell(P_r)$, where $F_\ell$ is the degree $(2\ell - 1)$ $\mathbb{Z}$-polynomial from Lemma 2.11 and we regard $P_r$ as a $\mathbb{Z}$-polynomial in the natural way. We can represent $Q_r$ as

$$Q_r = \sum_{S \subseteq [s] : |S| \le d_2} a_S \prod_{i \in S} g_i$$

in time $\mathrm{poly}(n)\binom{s}{d_2}$, where $d_2 = O(\ell \lg^9 n \lg^5(1/\varepsilon)\sqrt{m})$ and $a_S = n^{O(d_2)}$. For each $\prod_{i\in S} g_i \in \mathbf{AND}\circ\mathbf{SYM}$, we apply Lemma 2.5 and obtain a circuit $g_S \in \mathbf{SYM}$ with $n^{O(d_2)}$ wires. Let $Q'_r := \sum_{S\subseteq[s]:|S|\leq d} a_S g_S$. Finally we define $R : \{0,1\}^{n-n'} \to \mathbb{Z}$ as

$$R(y) := \sum_{a\in\{0,1\}^{n'}, r\in\{0,1\}^l} Q'_r(y,a) \bmod 2^\ell.$$

Note that if $2^\ell > 2^l$, then by Lemma 2.11 and the error probability of $Q'_r$, we have

$$C(x',a) = 1 \quad \Rightarrow \quad (1-\varepsilon)2^l \leq \left(\sum_{r\in\{0,1\}^l} Q'_r(y,a) \bmod 2^\ell\right) \leq 2^l,$$

$$C(x',a) = 0 \quad \Rightarrow \quad 0 \leq \left(\sum_{r\in\{0,1\}^l} Q'_r(y,a) \bmod 2^\ell\right) \leq \varepsilon 2^l.$$

In addition, if $2^\ell > 2^{n'}2^l$, then we have

$$R(y) \in (2^l(1-\varepsilon)K(y), 2^l\{(1-\varepsilon)K(y) + \varepsilon 2^{n'}\}).$$

If we set $\varepsilon < 1/2^{n'+1}$ and define $\tilde{R}(y)$ as the nearest integer of $R(y)/2^l$, then $\tilde{R}(y) = K(y)$ holds.

We set $n' = (n/(\sqrt{m}\lg^{c_1} n))^{c_2}$ for sufficiently large $c_1 > 0$ and small $c_2 > 0$, $\varepsilon = 1/2^{n'+2}$ and $\ell = n' + l + 1$. Then, we see that the construction of $R$ takes time at most $2^{n-n'}$. Furthermore, for each $i$, the $i$th bit of the binary representation of $R(y)$ can be represented as a circuit in $\mathbf{SYM}\circ\mathbf{SYM}$ so that the condition of Lemma 2.6 is satisfied as an $n-n'$-variate circuit.

This completes the proof. □

## 5 Circuit Lower Bounds

In this section, we give a proof sketch of Theorem 1.2.

We use the connection between satisfiability algorithms and circuit lower bounds due to Ben-Sasson and Viola [9]. Let $C_n$ be a set of functions from $\{0,1\}^n$ to $\{0,1\}$. $C_n$ is *closed under projections* if for all $f \in C_n$, indices $i, j \leq n$ and a bit $b$, it holds that

$$\neg f, f(x_1,\ldots,x_{i-1},x_j\oplus b,x_{i+1},\ldots,x_n), f(x_1,\ldots,x_{i-1},b,x_{i+1},\ldots,x_n) \in C_n.$$

$C_n$ is *efficiently closed under projections* if it is closed under projections and give a description of $f \in C_n$, we can compute in $\mathrm{poly}(|f|)$, descriptions of

$$\neg f, f(x_1,\ldots,x_{i-1},x_j\oplus b,x_{i+1},\ldots,x_n), f(x_1,\ldots,x_{i-1},b,x_{i+1},\ldots,x_n) \in C_n.$$

**Theorem 5.1** ([9]). *Let $C_n$ be efficiently closed under projections. If the satisfiability problem of the form $f_1 \wedge f_2 \wedge f_3$ for $f_1, f_2, f_3 \in C_{n+O(\lg n)}$ can be deterministically solved in time $2^{n-\omega(\lg n)}$, then there exists a language $L \in \mathrm{E}^{\mathrm{NP}}$ such that $L_n \notin C_n$ holds for infinitely many $n$. Here $L_n$ denotes the indicator function of $L \cap \{0,1\}^n$.*

It is easy to see that we can modify the proof of Theorem 4.1 to handle a circuit of the form $C_1 \wedge C_2 \wedge C_3$, where $C_1, C_2, C_3 \in (\mathbf{SYM}\cup\mathbf{THR})\circ(\mathbf{SYM}\cup\mathbf{THR})$, because the degree of the "final polynomial" is larger by a factor of at most 3. The class of depth 2 circuits with $m$ symmetric and linear threshold gates is clearly efficiently closed under projections. This completes the proof of Theorem 1.2.

## Acknowledgment

We are deeply grateful to Srikanth Srinivasan who explained us why pseudorandom generators due to [46, 24] are sufficient to obtain randomness efficient versions of his constructions from [58] and kindly allowed us to use the idea in this paper. We would like to thank Ruiwen Chen, Ramamohan Paturi, Rahul Santhanam and Stefan Schneider for useful discussion and valuable comments.

## References

[1] J. Alman and R. Williams. Probabilistic polynomials and Hamming nearest neighbors. In *Proceedings of the IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 136–150, 2015.

[2] K. Amano and A. Saito. A nonuniform circuit class with multilayer of threshold gates having super quasi polynomial size lower bounds against NEXP. In *Proceedings of the 9th International Conference on Language and Automata Theory and Applications (LATA)*, pages 461–472, 2015.

[3] K. Amano and A. Saito. A satisfiability algorithm for some class of dense depth two threshold circuits. *IEICE Transactions*, 98-D(1):108–118, 2015.

[4] P. Beame, R. Impagliazzo, and S. Srinivasan. Approximating $\mathrm{AC}^0$ by small height decision trees and a deterministic algorithm for $\#\mathrm{AC}^0$ SAT. In *Proceedings of the 27th Conference on Computational Complexity (CCC)*, pages 117–125, 2012.

[5] R. Beigel. The polynomial method in circuit complexity. In *Proceedings of the 8th Annual Structure in Complexity Theory Conference*, pages 82–95, 1993.

[6] R. Beigel. When do extra majority gates help? Polylog($n$) majority gates are equivalent to one. *Computational Complexity*, 4:314–324, 1994.

[7] R. Beigel, N. Reingold, and D. A. Spielman. The perceptron strikes back. In *Proceedings of the Sixth Annual Structure in Complexity Theory Conference*, pages 286–291, 1991.

[8] R. Beigel and J. Tarui. On ACC. *Computational Complexity*, 4:350–366, 1994.

[9] E. Ben-Sasson and E. Viola. Short PCPs with projection queries. In *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP), Part I*, pages 163–173, 2014.

[10] N. Bhatnagar, P. Gopalan, and R. J. Lipton. Symmetric polynomials over $\mathrm{Z}_m$ and simultaneous communication protocols. *J. Comput. Syst. Sci.*, 72(2):252–285, 2006.

[11] T. M. Chan and R. Williams. Deterministic APSP, orthogonal vectors, and more: Quickly derandomizing Razborov-Smolensky. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1246–1255, 2016.

[12] R. Chen. Satisfiability algorithms and lower bounds for Boolean formulas over finite bases. In *Proceedings of the 40th International Symposium on Mathematical Foundations of Computer Science (MFCS), Part II*, pages 223–234, 2015.

[13] R. Chen and V. Kabanets. Correlation bounds and #SAT algorithms for small linear-size circuits. In *Proceedings of the 21st International Conference on Computing and Combinatorics (COCOON)*, pages 211–222, 2015.

[14] R. Chen, V. Kabanets, A. Kolokolova, R. Shaltiel, and D. Zuckerman. Mining circuit lower bound proofs for meta-algorithms. *Computational Complexity*, 24(2):333–392, 2015.

[15] R. Chen, V. Kabanets, and N. Saurabh. An improved deterministic #SAT algorithm for small De Morgan formulas. In *Proceedings of the 39th International Symposium on Mathematical Foundations of Computer Science (MFCS), Part II*, pages 165–176, 2014.

[16] R. Chen and R. Santhanam. Improved algorithms for sparse MAX-SAT and MAX-*k*-CSP. In *Proceedings of the 18th International Conference on Theory and Applications of Satisfiability Testing (SAT)*, pages 33–45, 2015.

[17] R. Chen, R. Santhanam, and S. Srinivasan. Average-case lower bounds and satisfiability algorithms for small threshold circuits. In *Proceedings of the 31th Conference on Computational Complexity (CCC)*, pages 1:1–1:35, 2016.

[18] T. Y. Chow. Almost-natural proofs. *J. Comput. Syst. Sci.*, 77(4):728–737, 2011.

[19] J. Forster. A linear lower bound on the unbounded error probabilistic communication complexity. *J. Comput. Syst. Sci.*, 65(4):612–625, 2002.

[20] J. Forster, M. Krause, S. V. Lokam, R. Mubarakzjanov, N. Schmitt, and H. Simon. Relations between communication complexity, linear arrangements, and computational complexity. In *Proceedings of the 21st Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 171–182, 2001.

[21] M. Goldmann. On the power of a threshold gate at the top. *Inf. Process. Lett.*, 63(6):287–293, 1997.

[22] M. Goldmann, J. Håstad, and A. A. Razborov. Majority gates vs. general weighted threshold gates. *Computational Complexity*, 2:277–300, 1992.

[23] A. Golovnev, A. S. Kulikov, A. Smal, and S. Tamaki. Circuit size lower bounds and #SAT upper bounds through a general framework. In *Proceedings of the 41st International Symposium on Mathematical Foundations of Computer Science (MFCS)*, 2016, to appear.

[24] P. Gopalan, D. Kane, and R. Meka. Pseudorandomness via the discrete Fourier transform. In *Proceedings of the IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 903–922, 2015.

[25] P. Gopalan and R. A. Servedio. Learning and lower bounds for $AC^0$ with threshold gates. In *Proceedings of the 13th APPROX and the 14th RANDOM*, pages 588–601, 2010.

[26] H. D. Gröger and G. Turán. On linear decision trees computing Boolean functions. In *Proceedings of the 18th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 707–718, 1991.

[27] 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.

[28] J. Håstad and M. Goldmann. On the power of small-depth threshold circuits. *Computational Complexity*, 1:113–129, 1991.

[29] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[30] T. Hofmeister. A note on the simulation of exponential threshold weights. In *Proceedings of the Second Annual International Conference on Computing and Combinatorics (COCOON)*, pages 136–141, 1996.

[31] R. Impagliazzo, W. Matthews, and R. Paturi. A satisfiability algorithm for $\mathrm{AC}^0$. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 961–972, 2012.

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

[33] R. Impagliazzo, R. Paturi, and S. Schneider. A satisfiability algorithm for sparse depth two threshold circuits. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 479–488, 2013.

[34] H. Jahanjou, E. Miles, and E. Viola. Local reductions. In *Proceedings of the 42nd International Colloquium on Automata, Languages, and Programming (ICALP), Part I*, pages 749–760, 2015.

[35] D. M. Kane and R. Williams. Super-linear gate and super-quadratic wire lower bounds for depth-two and depth-three threshold circuits. In *Proceedings of the 48th ACM Symposium on Theory of Computing Conference (STOC)*, pages 633–643, 2016.

[36] M. Krause. Geometric arguments yield better bounds for threshold circuits and distributed computing. *Theor. Comput. Sci.*, 156(1&2):99–117, 1996.

[37] M. Krause and S. Lucks. Pseudorandom functions in $\mathrm{TC}^0$ and cryptographic limitations to proving lower bounds. *Computational Complexity*, 10(4):297–313, 2001.

[38] M. Krause and P. Pudlák. On the computational power of depth-2 circuits with threshold and modulo gates. *Theor. Comput. Sci.*, 174(1-2):137–156, 1997.

[39] M. Krause and S. Waack. Variation ranks of communication matrices and lower bounds for depth-two circuits having nearly symmetric gates with unbounded fan-in. *Mathematical Systems Theory*, 28(6):553–564, 1995.

[40] S. Lovett and S. Srinivasan. Correlation bounds for poly-size $\mathrm{AC}^0$ circuits with $n^{1-o(1)}$ symmetric gates. In *Proceedings of the 14th APPROX 2011 and the 15th RANDOM*, pages 640–651, 2011.

[41] A. Maciel and D. Thérien. Threshold circuits of small majority-depth. *Inf. Comput.*, 146(1):55–83, 1998.

[42] E. Miles and E. Viola. Substitution-permutation networks, pseudorandom functions, and natural proofs. *J. ACM*, 62(6):46, 2015.

[43] S. Muroga. *Threshold Logic and Its Applications*. John Wiley & Sons, 1971.

[44] A. Nagao, K. Seto, and J. Teruyama. A moderately exponential time algorithm for $k$-IBDD satisfiability. In *Proceedings of the 14th International Symposium, on Algorithms and Data Structures (WADS)*, pages 554–565, 2015.

[45] M. Naor and O. Reingold. Number-theoretic constructions of efficient pseudo-random functions. *J. ACM*, 51(2):231–262, 2004.

[46] N. Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12(4):449–461, 1992.

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

[48] S. Nurk. An $O(2^{0.4058m})$ upper bound for circuit SAT, 2009.

[49] I. C. Oliveira. Algorithms versus circuit lower bounds. *Electronic Colloquium on Computational Complexity (ECCC)*, TR13-117, 2013.

[50] R. Paturi and M. E. Saks. Approximating threshold circuits by rational functions. *Inf. Comput.*, 112(2):257–272, 1994.

[51] V. V. Podolskii. Exponential lower bound for bounded depth circuits with few threshold gates. *Inf. Process. Lett.*, 112(7):267–271, 2012.

[52] A. A. Razborov and S. Rudich. Natural proofs. *J. Comput. Syst. Sci.*, 55(1):24–35, 1997.

[53] A. A. Razborov and A. Wigderson. $n^{\Omega(\log n)}$ lower bounds on the size of depth-3 threshold circuits with AND gates at the bottom. *Inf. Process. Lett.*, 45(6):303–307, 1993.

[54] T. Sakai, K. Seto, S. Tamaki, and J. Teruyama. Bounded depth circuits with weighted symmetric gates: Satisfiability, lower bounds and compression. In *Proceedings of the 41st International Symposium on Mathematical Foundations of Computer Science (MFCS)*, 2016, to appear.

[55] R. Santhanam. Fighting perebor: New and improved algorithms for formula and QBF satisfiability. In *Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 183–192, 2010.

[56] R. Santhanam. Ironic complicity: Satisfiability algorithms and circuit lower bounds. *Bulletin of the EATCS*, 106:31–52, 2012.

[57] K. Seto and S. Tamaki. A satisfiability algorithm and average-case hardness for formulas over the full binary basis. *Computational Complexity*, 22(2):245–274, 2013.

[58] S. Srinivasan. On improved degree lower bounds for polynomial approximation. In *Proceedings of the 33rd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 201–212, 2013.

[59] A. Tal. #SAT algorithms from shrinkage. *Electronic Colloquium on Computational Complexity (ECCC)*, TR15-114, 2015.

[60] J. Tarui. Probablistic polynomials, $\mathrm{AC}^0$ functions, and the polynomial-time hierarchy. *Theor. Comput. Sci.*, 113(1):167–183, 1993.

[61] S. Toda. PP is as hard as the polynomial-time hierarchy. *SIAM J. Comput.*, 20(5):865–877, 1991.

[62] F. Wang. NEXP does not have non-uniform quasipolynomial-size ACC circuits of $o(\log\log n)$ depth. In *Proceedings of the 8th Annual Conference on Theory and Applications of Models of Computation (TAMC)*, pages 164–170, 2011.

[63] R. Williams. Improving exhaustive search implies superpolynomial lower bounds. *SIAM J. Comput.*, 42(3):1218–1244, 2013.

[64] R. Williams. Natural proofs versus derandomization. In *Proceedings of the 45th ACM Symposium on Theory of Computing Conference (STOC)*, pages 21–30, 2013.

[65] R. Williams. Algorithms for circuits and circuits for algorithms. In *Proceedings of the 29th Annual IEEE Conference on Computational Complexity (CCC)*, pages 248–261, 2014.

[66] R. Williams. New algorithms and lower bounds for circuits with linear threshold gates. In *Proceedings of the 46th Symposium on Theory of Computing (STOC)*, pages 194–202, 2014. Also arXiv:1401.2444.

[67] R. Williams. Nonuniform ACC circuit lower bounds. *J. ACM*, 61(1):2, 2014.

[68] R. Williams. The polynomial method in circuit complexity applied to algorithm design (invited talk). In *Proceedings of the 34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 47–60, 2014.

[69] A. C. Yao. On ACC and threshold circuits. In *Proceedings of the 31st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 619–627, 1990.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de