On the Size of Depth-Two Threshold Circuits for the Inner Product Mod 2 Function

Kazuyuki Amano (✉)

Department of Computer Science, Gunma University,
1-5-1 Tenjin, Kiryu, Gunma 376-8515, Japan
amano@gunma-u.ac.jp

Abstract. In this paper, we study the size of depth-two threshold circuits computing the inner product mod 2 function $\text{IP2}_n(x_1, \dots, x_n, y_1, \dots, y_n) := \sum_i x_i y_i \pmod{2}$. First, we reveal that IP2_n can be computed by a depth-two threshold circuit of size significantly smaller than a folklore construction of size $O(2^n)$. Namely, we give a construction of such a circuit (denoted by $\text{THR} \circ \text{THR}$ circuit) of size $O(1.682^n)$. We also give an upper bound of $O(1.899^n)$ for the case that the weights of the top threshold gate are polynomially bounded (denoted by $\text{MAJ} \circ \text{THR}$ circuit). Second, we give new lower bounds on the size of depth-two circuits of some special form; the top gate is an unbounded weight threshold gate and the bottom gates are symmetric gates (denoted by $\text{THR} \circ \text{SYM}$ circuit). We show that any such circuit computing IP2_n has size $\Omega((1.5 - \epsilon)^n)$ for every constant $\epsilon > 0$. This improves the previous bound of $\Omega(\sqrt{2}^n/n)$ based on the sign-rank method due to Forster et al. [JCSS '02, FSTTCS '01]. Our technique has a unique feature that the lower bound is obtained by giving an explicit feasible solution to (the dual of) a certain linear programming problem. In fact, the problem itself was presented by the author over a decade ago [MFCS '05], and finding a good solution is an actual contribution of this work.

Keywords: Circuit complexity · Threshold circuits · Linear programming · Upper bounds · Lower bounds

1 Introduction

The problem of proving strong lower bounds on the size (i.e., the number of gates) of depth-two threshold circuits computing an explicit Boolean function is a big challenge in complexity theory. Currently, we cannot refute that every function in the class NEXP (non-deterministic exponential time) can be computed by a polynomial-size depth-two circuit consisting of threshold gates with unbounded weights (denoted by $\text{THR} \circ \text{THR}$ circuit). There is a long line of research aiming for understanding the computational power and the limitation of depth-two threshold circuits (e.g. [5, 9, 10, 13, 14] or see an excellent book [12, Chapter 11.10]). The strongest known lower bound on the size of $\text{THR} \circ \text{THR}$ circuits for a function in NP is $\Omega(n^{3/2})$ due to Kane and Williams [13].

In this paper, we focus on the size complexity of depth-two threshold circuits for the *inner product mod 2* function:

$$\text{IP2}_n(x_1, \dots, x_n, y_1, \dots, y_n) = \sum_{i=1}^n (x_i \wedge y_i) \pmod{2}.$$

The inner product mod 2 function IP2_n has been widely studied in the context of depth-two threshold circuits (e.g., [7, 10, 13]).

It is a long standing open question whether IP2_n has a polynomial size depth-two threshold circuit with unbounded weights threshold gates in both layers. If we restrict the weights of threshold gates in one of two layers to be polynomial, then strong lower bounds are known. Let MAJ denote the class of threshold functions whose weights are bounded to be $\mathbb{Z} \cap [-\text{poly}(n), \text{poly}(n)]$. Hajnal et al. [10] proved that every $\text{MAJ} \circ \text{THR}$ circuit computing IP2_n has size $\Omega(2^{(1/3-\epsilon)n})$ using the discriminator method. An exponential lower bound were also shown by Nisan [16] using a communication complexity argument. Forster et al. [7, 8] proved that every $\text{THR} \circ \text{MAJ}$ circuit computing IP2_n has size $\Omega(\sqrt{2}^n / \text{poly}(n))$ by lowerbounding the sign-rank of the communication matrix of IP2_n .

Note that IP2_n has an $O(n)$ size threshold circuit of *depth-three*; in the first layer, we use n gates to compute $x_i \wedge y_i$ for each i , and then in the second and third layer, we use $O(n)$ gates to compute the parity of the outputs of them. If the gates at the bottom layer are restricted to be And, Exclusive-or or Symmetric gates, stronger lower bounds for IP2_n are known (see Table 1). Remark that, in recent years, several results providing the separation between depth-two and depth-three threshold circuits were given for *real-valued* functions (e.g., [6, 18]). However, to the best of our knowledge, the arguments used in these works can not directly be applied for Boolean functions.

Table 1. Known upper and lower bounds on the size of depth-two circuits using threshold gates that computes IP2_n . Entries marked with (*) are shown in this paper. Unmarked results are folklore.

Circuit type	Lower bound	Upper bound
$\text{THR} \circ \text{AND}$	2^n [3]	2^n
$\text{THR} \circ \text{XOR}$	2^n [4]	$O(2.966^n)$ [2, 19]
$\text{THR} \circ \text{SYM}$	$\Omega((1.5 - \epsilon)^n)$ (*)	2^n
$\text{THR} \circ \text{MAJ}$	$\Omega((\sqrt{2}^n / \text{poly}(n)))$ [7, 8]	2^n
$\text{THR} \circ \text{THR}$	$\Omega(n)$ [17]	$O(1.682^n)$ (*)
$\text{MAJ} \circ \text{THR}$	$\Omega(2^{(1/3-\epsilon)n})$ [10]	$O(1.899^n)$ (*)

1.1 Our Contributions

The contribution of this work is twofold.

First, we consider *upper bounds* on the size of depth-two threshold circuits for IP2_n . Although we know that lower bounds are more preferable, we pursue upper bounds because we think that the lack of knowledge on good upper bounds for the problem is one of the reasons why we could not obtain a good lower bound.

It is folklore that IP2_n can be computed by a $\text{THR} \circ \text{AND}$ circuit (hence also by a $\text{THR} \circ \text{THR}$ circuit) of size 2^n by applying the inclusion-exclusion formula. Namely,

$$\text{IP2}_n(x_1, \dots, x_n, y_1, \dots, y_n) = \sum_{\emptyset \neq S \subseteq \{1, \dots, n\}} (-2)^{|S|-1} \prod_{i \in S} x_i y_i.$$

To the best of our knowledge, no asymptotically better bound has not been published. Note that IP2_n has $2n$ input variables and the construction via the DNF representation of IP2_n needs $\sim 3^n$ gates.

In this work, we show that IP2_n has a depth-two threshold circuit of size significantly smaller than 2^n . Namely, we give an explicit construction of a $\text{THR} \circ \text{THR}$ circuit of size $O(1.682^n)$ as well as a $\text{MAJ} \circ \text{THR}$ circuit of size $O(1.899^n)$ computing IP2_n .

The second contribution of this work is to give a new lower bound on the size of depth-two threshold circuits with some special restriction on the bottom gates. A *symmetric gate* is a gate that takes Boolean inputs whose output is depending only on the number of one's in inputs. Let $\text{THR} \circ \text{SYM}$ denote depth-two circuits consisting of a threshold gate with unbounded weights at the top and symmetric gates at the bottom.

In [7], Forster established a breakthrough result that the sign-rank of the $2^n \times 2^n$ Hadamard matrix is $\Omega(\sqrt{2}^n)$. Here the sign-rank of a matrix $M = (M_{i,j})$ with nonzero entries is the least rank of a matrix $A = (A_{i,j})$ with $M_{i,j}A_{i,j} > 0$ for all i and j . By combining this result and a simple fact that the communication matrix of any symmetric function has rank at most $n + 1$, Forster et al. [8] established an $\Omega(\sqrt{2}^n/n)$ lower bound on the size of $\text{THR} \circ \text{SYM}$ circuits for IP2_n .

In this paper, we improve their bound to $\Omega((1.5 - \epsilon)^n)$. Although the improvement is somewhat limited, our method has a unique feature; the lower bound is obtained by giving an explicit feasible solution to a certain linear programming problem.

Over a decade ago, building on the work of Basu et al. [3], the author developed an LP-based method to obtain a lower bound on the size of $\text{THR} \circ \text{SYM}$ circuits for IP2_n [1]. In [1], we showed that the problem of obtaining a lower bound on the size of such circuits can be reduced to the problem of solving a certain linear programming problem. Then we solved an obtained linear programming problem over 2^{16} variables using an LP solver to establish a lower bound of $\Omega(1.3638^n)$ on the size of $\text{THR} \circ \text{SYM}$ circuits for IP2_n . However, the problem of determining a highest lower bound that can be obtained by our LP-based method was left as an open problem in [1].

In this work, we show that this limit is in fact $\Omega((1.5 - \epsilon)^n)$, surpassing the $\Omega(\sqrt{2}^n/n)$ bounds obtained by the sign-rank method. We achieve this by giving an explicit feasible solution to the *dual* of the linear programming problem presented in [1] and estimating the value of the objective function. Showing this is an actual contribution of the second part of this work.

The rest of the paper is organized as follows. In Sect. 2, we introduce some notations. In Sect. 3, we give new upper bounds on the size of depth-two threshold circuits for IP2_n . Then in Sect. 4, we review an LP-based lower bounds method presented in our previous work [1], and establish a new lower bound on the size of $\text{THR} \circ \text{SYM}$ circuits for IP2_n .

2 Preliminaries

For an integer $n \geq 1$, $[n]$ denotes the set $\{1, 2, \dots, n\}$. The *inner product mod 2 function* IP2_n is a Boolean function over $2n$ variables defined by

$$\text{IP2}_n(x_1, \dots, x_n, y_1, \dots, y_n) = \sum_{i=1}^n (x_i \wedge y_i) \pmod{2}.$$

For a Boolean predicate P , let $\llbracket P \rrbracket$ denote the Iverson bracket function defined as $\llbracket P \rrbracket = 1$ if P is true and $\llbracket P \rrbracket = 0$ if P is false.

Let $x_1, \dots, x_n \in \{0, 1\}$ be Boolean variables. A *linear threshold function* is a Boolean function of the form

$$\llbracket w_1 x_1 + \dots + w_n x_n \geq t \rrbracket,$$

for some $w_1, \dots, w_n, t \in \mathbb{R}$. Similarly, an *exact threshold function* is a Boolean function of the form

$$\llbracket w_1 x_1 + \dots + w_n x_n = t \rrbracket.$$

We call $w_1, \dots, w_n$ the *weights* and t the *threshold*. It is well known that, without loss of generality, we can assume that the weights and the threshold are integers of absolute value $2^{O(n \log n)}$ [15]. Hence, hereafter, we assume that the weights and the threshold are all integers. A gate that computes a linear threshold function is called a *threshold gate*. The class of all linear threshold functions (exact threshold functions, respectively) is denoted by THR (ETHR , respectively).

As usual, a depth-two circuit such that the top gate computes a function in C , and every bottom gate computes a function in $\mathcal{D}$ is called a $C \circ \mathcal{D}$ circuit. For example, a $\text{THR} \circ \text{THR}$ circuit is a depth-two circuit with threshold gates of unbounded weights in both layers. The *size* of a depth-two circuit is defined to be the number of gates in the bottom layer. The *size complexity* of a Boolean function f for $C \circ \mathcal{D}$ circuits is the minimum size of a $C \circ \mathcal{D}$ circuit computing f .

A *majority gate* is a gate computing a linear threshold function with additional restriction that $w_i \in \{-1, 1\}$ for all i . Here the threshold t can be an arbitrary value, i.e., is not restricted to be the half of the number of input variables. The class of functions computed by a majority gate is denoted by MAJ . In our definition, a majority gate is allowed to read a variable multiple times. For example, we can say that the function

$$\llbracket x_1 - 2x_2 + 3x_3 \geq 2 \rrbracket$$

is computed by a majority gate of fan-in $1 + 2 + 3 = 6$. Remark that a majority gate is often defined as a gate that computes a linear threshold function with polynomially

bounded weights. If we adapt this definition of majority gates, the size complexity may be reduced by at most a polynomial factor. However, such a difference will not affect all the results described in this paper.

A function $f : \{0, 1\}^n \rightarrow \mathbb{R}$ is called *symmetric* if the value of f depends only on the number of ones in the input. A gate that computes a symmetric function is called a *symmetric gate* and the class of all symmetric functions is denoted by **SYM**. Note that a symmetric gate is usually defined as a Boolean gate, i.e., it outputs a binary value. In this paper, we extend the domain from $\{0, 1\}$ to $\mathbb{R}$. By this extension, the set of symmetric functions turns out to be closed under linear combinations. This property is useful when we view a threshold-of-symmetric circuit as (the sign of) a real polynomial (see Sect. 4.1). Note also that a symmetric gate can simulate all of AND, OR, the modulo gate. It can also simulate a restrict version of the majority gate where the gate reads each variable at most once and all the weights are restricted to be 1.

3 Upper Bounds

In this section, we give upper bounds on the size of depth-two threshold circuits for IP2_n , which is significantly smaller than a folklore bound of $O(2^n)$.

We begin with two simple lemmas about exact threshold functions. Both lemmas were appeared in [11].

Lemma 1 [11]. *Suppose that a Boolean function f can be computed by a $\text{THR} \circ \text{ETHR}$ circuit of size s . Then, f can be computed by a $\text{THR} \circ \text{THR}$ circuit of size at most $2s$. The same relationship holds for $\text{MAJ} \circ \text{ETHR}$ and $\text{MAJ} \circ \text{THR}$ circuits.*

Lemma 2 [11]. *The AND of an arbitrary number of exact threshold functions is also an exact threshold function. In other words, the class of exact threshold functions is closed under the AND operation.*

Before stating our upper bounds, we describe an idea of our construction. Consider the function $\text{IP2}_2(x_1, x_2, y_1, y_2)$. Define two exact threshold functions g_1 and g_2 as follows.

$$\begin{aligned} g_1(x_1, x_2, y_1, y_2) &= \llbracket x_1 + x_2 + y_1 + y_2 = 1 \rrbracket, \\ g_2(x_1, x_2, y_1, y_2) &= \llbracket x_1 - x_2 + y_1 - y_2 = 0 \rrbracket. \end{aligned}$$

It is easy to verify that

$$\text{IP2}_2(x_1, x_2, y_1, y_2) = \text{sgn}(2 \cdot g_1(x_1, x_2, y_1, y_2) + 2 \cdot g_2(x_1, x_2, y_1, y_2) - 1),$$

where $\text{sgn}(v)$ is defined to be 0 if $v > 0$ and is 1 if $v < 0$.

Then, when n is even, $\text{IP2}_n(x_1, \dots, x_n, y_1, \dots, y_n)$ is given by

$$\text{sgn} \left(\prod_{i \in [\frac{n}{2}]} (2 \cdot g_1(x_{2i-1}, x_{2i}, y_{2i-1}, y_{2i}) + 2 \cdot g_2(x_{2i-1}, x_{2i}, y_{2i-1}, y_{2i}) - 1) \right). \quad (1)$$

By expanding the product in Eq. (1), we can obtain a polynomial of $3^{n/2}$ terms in which each term is an AND of exact threshold functions. By Lemma 2, we can express each term by a single ETHR gate. Therefore, we have a $\text{THR} \circ \text{ETHR}$ circuit of size $O(3^{n/2}) = O(1.733^n)$ for IP2_n , and also have a $\text{THR} \circ \text{THR}$ circuit of the same order by Lemma 1.

It is natural to expect that we can obtain a better bound by considering IP2_k for $k > 2$ as a base case. These ideas can be summarized as the following theorem.

Theorem 3. *Let k be a positive integer. We write $\mathbf{x} = (x_1, \dots, x_k) \in \{0, 1\}^k$ and $\mathbf{y} = (y_1, \dots, y_k) \in \{0, 1\}^k$. Suppose that IP2_k can be represented by the sign of the linear combination of ℓ exact threshold functions where all weights are integers, i.e.,*

$$\text{IP2}_k(\mathbf{x}, \mathbf{y}) = \text{sgn} \left(\sum_{i \in [\ell]} w_i C_i(\mathbf{x}, \mathbf{y}) \right),$$

where $w_i \in \mathbb{Z}$ and $C_i \in \text{ETHR}$ for $i \in [\ell]$. Then,

1. The size complexity of IP2_n for $\text{THR} \circ \text{ETHR}$ circuits as well as $\text{THR} \circ \text{THR}$ circuits is $O((\ell^{1/k})^n)$,
2. The size complexity of IP2_n for $\text{MAJ} \circ \text{THR}$ circuits is $O((\sum_{i \in [\ell]} |w_i|)^{n/k})$.

Proof (Sketch). First, observe that IP2_n is just a PARITY of n/k copies of IP2_k . Replace each IP2_k with a constructed ℓ -gate $\text{THR} \circ \text{ETHR}$ circuit. The PARITY of n/k THR of ℓ ETHRs can be written as the sign of the product of n/k sums of ℓ ETHRs. Applying distributivity to the product of sums, we get a sum of $\ell^{n/k}$ products of ETHRs. But the product of a bunch of ETHRs can be written as one ETHR, so we get a THR of $\ell^{n/k}$ ETHRs, completing the proof of Statement 1 of the theorem. The proof for Statement 2 is similar. $\square$

With the aid of computers, we found a formula of length 8 for IP2_4 as well as a formula of total weight 13 for IP2_4 that lead us to the following theorems.

Theorem 4. *The size complexity of IP2_n for $\text{THR} \circ \text{ETHR}$ circuits (and also for $\text{THR} \circ \text{THR}$ circuits) is $O(8^{n/4}) = O(1.682^n)$.*

Theorem 5. *The size complexity of IP2_n for $\text{MAJ} \circ \text{THR}$ circuits is $O(13^{n/4}) = O(1.899^n)$.*

Proof of Theorem 4. Let $\{x_1, \dots, x_4, y_1, \dots, y_4\}$ denote the input variables for IP2_4 . For $i \in [4]$, we write $z_i := x_i + y_i$. We introduce the following seven exact threshold functions and write them as $g_1, \dots, g_7$.

$$\begin{aligned} \llbracket -z_1 + z_2 + z_3 + z_4 = 1 \rrbracket, & \quad \llbracket z_1 - z_2 + z_3 + z_4 = 1 \rrbracket, \\ \llbracket z_1 + z_2 - z_3 + z_4 = 1 \rrbracket, & \quad \llbracket z_1 + z_2 + z_3 - z_4 = 1 \rrbracket, \\ \llbracket z_1 - z_2 - z_3 + z_4 = 0 \rrbracket, & \quad \llbracket z_1 - z_2 + z_3 - z_4 = 0 \rrbracket, \\ \llbracket z_1 + z_2 - z_3 - z_4 = 0 \rrbracket. & \end{aligned}$$

It is elementary to verify that

$$\text{IP2}_4(x_1, \dots, x_4, y_1, \dots, y_4) = \text{sgn} \left(-3 + 2 \sum_{i \in [7]} g_i(z_1, z_2, z_3, z_4) \right).$$

This gives a desired bound by Theorem 3. $\square$

Proof of Theorem 5. Let $\{x_1, \dots, x_4, y_1, \dots, y_4\}$ denote the input variables for IP2_4 . For $i \in [4]$, we write $z_i := x_i + y_i$. We introduce the following twelve exact threshold functions and write them as $g_1, \dots, g_{12}$.

$$\begin{aligned} \llbracket 3z_1 - 3z_2 + 2z_3 + 4z_4 = 8 \rrbracket, & \quad \llbracket 3z_1 - 3z_2 + 4z_3 + 2z_4 = 8 \rrbracket, \\ \llbracket -3z_1 + 3z_2 + 2z_3 + 4z_4 = 8 \rrbracket, & \quad \llbracket -3z_1 + 3z_2 + 4z_3 + 2z_4 = 8 \rrbracket, \\ \llbracket 2z_1 + 4z_2 + 3z_3 - 3z_4 = 8 \rrbracket, & \quad \llbracket 4z_1 + 2z_2 + 3z_3 - 3z_4 = 8 \rrbracket, \\ \llbracket 2z_1 + 4z_2 - 3z_3 + 3z_4 = 8 \rrbracket, & \quad \llbracket 4z_1 + 2z_2 - 3z_3 + 3z_4 = 8 \rrbracket, \\ \llbracket 3z_1 + 3z_2 + 2z_3 + 4z_4 = 11 \rrbracket, & \quad \llbracket 3z_1 + 3z_2 + 4z_3 + 2z_4 = 11 \rrbracket, \\ \llbracket 2z_1 + 4z_2 + 3z_3 + 3z_4 = 11 \rrbracket. & \quad \llbracket 4z_1 + 2z_2 + 3z_3 + 3z_4 = 11 \rrbracket. \end{aligned}$$

It is elementary to verify that

$$\text{IP2}_4(x_1, \dots, x_4, y_1, \dots, y_4) = \text{sgn} \left(-1 + \sum_{i \in [12]} g_i(z_1, z_2, z_3, z_4) \right).$$

This gives a desired bound by Theorem 3. $\square$

It is plausible that our bounds can further be improved by considering IP2_k for $k \geq 5$ as a base case. We remark that, for the case of $\text{MAJ} \circ \text{THR}$ circuits, the following argument says that there is a barrier at $O(\sqrt{2}^n)$: The proof of Theorem 5 actually gives a construction of $\text{MAJ} \circ \text{ETHR}$ circuits for IP2_n . By applying the “discriminator lemma” developed in [10] carefully, we can prove an $\Omega(2^{(1/2-\epsilon)n})$ lower bound on the size complexity of IP2_n for $\text{MAJ} \circ \text{ETHR}$ circuits. Currently, we do not know such a barrier for $\text{THR} \circ \text{THR}$ circuits.

4 Lower Bounds for $\text{THR} \circ \text{SYM}$ Circuits

In this section, we show $\Omega((1.5 - \epsilon)^n)$ lower bounds on the size of depth-two circuits for IP2_n where the top gate is a threshold gate and the bottom gates are symmetric gates. In Sect. 4.1, we review our LP-based method presented in our previous work [1], and then we establish the lower bound in Sect. 4.2.

Throughout this section, we label the input variables of IP2_n as $\{x_1, \dots, x_{2n}\}$ and define $\text{IP2}_n(x_1, \dots, x_{2n}) := \sum_{i \in [n]} x_{2i-1} x_{2i} \pmod{2}$. This indexing is different from the one used in the previous section, but will be convenient for a later discussion.

4.1 LP-Based Method for Lower Bounds on Circuit Size

As defined before, we call a depth-two circuit with unbounded weights threshold gate at the top and symmetric gates at the bottom as a $\text{THR} \circ \text{SYM}$ circuit. For a Boolean function f , the size complexity of IP2_n for $\text{THR} \circ \text{SYM}$ circuits is simply denoted by $s(f)$. Throughout of this section, we treat a $\text{THR} \circ \text{SYM}$ circuit as the sign of a real polynomial.

Definition 6. We say that a real polynomial $P(x_1, \dots, x_n)$ sign represents a Boolean function f on n variables if, for every $(x_1, \dots, x_n) \in \{0, 1\}^n$,

$$\begin{aligned} f(x_1, \dots, x_n) = 0 &\implies P(x_1, \dots, x_n) > 0, \\ f(x_1, \dots, x_n) = 1 &\implies P(x_1, \dots, x_n) < 0. \end{aligned}$$

□

We consider a polynomial $P : \{0, 1\}^X \rightarrow \mathbb{R}$

$$P(X) = \sum_{S \subseteq X} w_S h_S(X), \quad (2)$$

where $w_S \in \mathbb{R}$ and h_S is a symmetric function over the set of variables S . The *support* of P is defined by $\text{supp}(P) := \{S \subseteq X \mid w_S \neq 0\}$. Obviously, $s(f)$ is equal to the minimum size of the support of a polynomial P of the form (2) that sign represents f .

A point of our method is to define the parameter z_k , which gives a lower bound on $s(f)$, by introducing a certain linear programming problem.

Recall that the input variables of $\text{IP}2_n$ is $X := \{x_1, \dots, x_{2n}\}$.

Let $z_0 = z_1 = 1$. For $k \geq 2$, the parameter z_k is defined inductively (on k) such that z_k is the minimum value of the objective function of the following linear programming problem. Let $X_k = \{x_1, x_2, \dots, x_{2k}\}$ be the first $2k$ variables of X . The program has 2^{2k} real-valued variables $\{q_T\}_{T \subseteq X_k}$ and $2k + 4\binom{k}{2}$ constraints.

$$\begin{aligned} &\text{Minimize} && \sum_{T \subseteq X_k} q_T, \\ &\text{Subject to} && \sum_{T: v \in T} q_T \geq z_{k-1} \quad (v \in X_k), \\ & && \sum_{T: \{|u,v\} \cap T| = 1} q_T \geq z_{k-2} \left(\begin{array}{l} i, j \in [k], i \neq j \\ u \in \{x_{2i-1}, x_{2i}\}, v \in \{x_{2j-1}, x_{2j}\} \end{array} \right), \\ & && q_T \geq 0 \quad (T \subseteq X_k). \end{aligned} \quad (3)$$

The key observation is the following.

Fact 7 ([1]). Suppose that $k \geq 2$. Let z_{k-1} and z_{k-2} be real numbers such that $s(\text{IP}2_n) \geq z_{k-1} \cdot s(\text{IP}2_{n-(k-1)})$ and $s(\text{IP}2_n) \geq z_{k-2} \cdot s(\text{IP}2_{n-(k-2)})$ for every n . Let z_k be the minimum value of the objective function of the LP problem (3). Then $s(\text{IP}2_n) \geq z_k \cdot s(\text{IP}2_{n-k})$.

The following corollary is immediate from Fact 7.

Corollary 8 ([1]). For every $k \geq 1$, $s(\text{IP}2_n) \geq (z_k^{1/k})^n$.

□

In the following, we give a sketch of the proof of Fact 7 for completeness.

Let $f : \{0, 1\}^X \rightarrow \mathbb{R}$ be a real function and $\rho : X \rightarrow \{0, 1, *\}$ be a partial assignment to X . Let $\text{res}(\rho)$ denote the set of variables that assigned a constant by ρ , i.e., $\text{res}(\rho) := \{v \in X \mid \rho(v) \neq *\}$. The restriction of f by ρ , denoted by $f|_\rho$, is the function obtained by setting x_i to $\rho(x_i)$ if $x_i \in \text{res}(\rho)$ and leaving x_i as a variable otherwise.

The restriction of a polynomial P of the form (2), denoted by $P|_\rho$, is defined similarly. First, replace each h_S in P by $h_S|_\rho$, which is a symmetric function over the set of

variables $S \setminus \text{res}(\rho)$. Then, if there are two (or more) functions $h_{S_1}|_\rho$ and $h_{S_2}|_\rho$ such that $S_1 \setminus \text{res}(\rho) = S_2 \setminus \text{res}(\rho)$, then they are merged into a single symmetric function. This is possible by the fact that the linear combination of two (or more) symmetric functions over the same set of variables is also a symmetric function.

For a polynomial P of the form (2), we decompose P into P_T 's for $T \subseteq X_k$ in such a way that

$$P_T(X) := \sum_{\substack{S \in \text{supp}(P) \\ S \cap X_k = T}} w_S h_S(X).$$

Let $\tilde{q}_T$ be the number of terms in P_T . Note that

$$P(X) = \sum_{T \subseteq X_k} P_T(X),$$

and

$$|\text{supp}(P)| = \sum_{T \subseteq X_k} \tilde{q}_T.$$

We use the following fact that is easy to verify but useful.

Fact 9 ([1]). *Let ρ_1 and ρ_2 be two partial assignments such that $\text{res}(\rho_1) = \text{res}(\rho_2)$. Then, $\sum_{v \in T \cap \text{res}(\rho_1)} \rho_1(v) = \sum_{v \in T \cap \text{res}(\rho_2)} \rho_2(v)$ implies $P_T|_{\rho_1} - P_T|_{\rho_2} \equiv 0$.* $\square$

Proof of Fact 7 (sketch). Suppose that a polynomial P of the form (2) sign-represents IP2_n . In what follows, we consider two types of pairs of partial assignments.

Type 1. Choose $i \in [k]$ and then choose $u \in \{x_{2i-1}, x_{2i}\}$. The unchosen variable in $\{x_{2i-1}, x_{2i}\}$ is denoted by v . Let ρ_1 and ρ_2 be two partial assignments such that $\text{res}(\rho_1) = \text{res}(\rho_2) = \{x_{2i-1}, x_{2i}\}$, $(\rho_1(v), \rho_1(u)) = (0, 1)$ and $(\rho_2(v), \rho_2(u)) = (1, 1)$.

A key observation is that for every such pair of partial assignments (ρ_1, ρ_2) , we have $\text{IP2}_n|_{\rho_1} \equiv \text{IP2}_{n-1}$ and $\text{IP2}_n|_{\rho_2} \equiv \overline{\text{IP2}_{n-1}}$. This implies that the polynomial $P|_{\rho_1} - P|_{\rho_2}$ sign-represents IP2_{n-1} . Fact 9 says that $P_T|_{\rho_1} - P_T|_{\rho_2}$ is vanished if $v \notin T$. Hence, we have

$$\sum_{T: v \in T} \tilde{q}_T \geq |\text{supp}(P|_{\rho_1} - P|_{\rho_2})| \geq s(\text{IP2}_{n-1}) \geq z_{k-1} \cdot s(\text{IP2}_{n-k}), \quad (4)$$

where the last inequality follows from the assumption in the statement of Fact 7. Let $q_T := \tilde{q}_T / s(\text{IP2}_{n-k})$ for $T \subseteq X_k$. By dividing both side of (4) by $s(\text{IP2}_{n-k})$, we have

$$\sum_{T: v \in T} q_T \geq z_{k-1},$$

which is the first constraint in the LP problem (3).

We also consider another type of partial assignments.

Type 2. Choose $i, j \in [k]$ such that $i \neq j$, and then choose $v \in \{x_{2i-1}, x_{2i}\}$ and $u \in \{x_{2j-1}, x_{2j}\}$. Let v' and u' be the unchosen variables in $\{x_{2i-1}, x_{2i}\}$ and $\{x_{2j-1}, x_{2j}\}$, respectively. Let ρ_1 and ρ_2 be two partial assignments such that

$\text{res}(\rho_1) = \text{res}(\rho_2) = \{x_{2i-1}, x_{2i}, x_{2j-1}, x_{2j}\}$, $(\rho_1(v), \rho_1(v'), \rho_1(u), \rho_1(u')) = (0, 1, 1, 0)$ and $(\rho_2(v), \rho_2(v'), \rho_2(u), \rho_2(u')) = (1, 1, 0, 0)$.

Similar to the case of Type 1, we have $\text{IP2}_n|_{\rho_1} \equiv \text{IP2}_{n-2}$ and $\text{IP2}_n|_{\rho_2} \equiv \overline{\text{IP2}_{n-2}}$, and hence $P|_{\rho_1} - P|_{\rho_2}$ sign represents IP2_{n-2} . In addition, $P_T|_{\rho_1} - P_T|_{\rho_2}$ is vanished if $|T \cap \{u, v\}|$ is zero or two. Hence, we have

$$\sum_{T: |T \cap \{u, v\}|=1} \tilde{q}_T \geq |\text{supp}(P|_{\rho_1} - P|_{\rho_2})| \geq s(\text{IP2}_{n-2}) \geq z_{k-2} \cdot s(\text{IP2}_{n-k}), \quad (5)$$

where the last inequality follows from the assumption of the statement in Fact 7. This inequality is equivalent to

$$\sum_{T: |T \cap \{u, v\}|=1} q_T \geq z_{k-2},$$

which is the second constraint in the LP problem (3).

If P is an optimal polynomial for IP2_n , then

$$s(\text{IP2}_n) = \sum_{T \subseteq X_k} \tilde{q}_T,$$

which is equivalent to

$$s(\text{IP2}_n) = \sum_{T \subseteq X_k} q_T \cdot s(\text{IP2}_{n-k}).$$

Therefore, the minimum value z_k of the objective function of the LP program (3) satisfies $s(\text{IP2}_n) \geq z_k \cdot s(\text{IP2}_{n-k})$. This completes the proof of Fact 7. $\square$

The LP problem (3) can easily be generated and solved by using a computer when k is small. In our previous work [1], we have succeeded to solve these problems by an LP solver for $k \leq 8$ (see Table 2). During this work, we could extend the table up to $k = 10$. The best lower bound obtained in this way is $\Omega(1.3808^n)$, but still weaker than a bound of $s(\text{IP2}_n) = \Omega(\sqrt{2}^n/n)$ due to Forster et al. [7, 8].

Obviously, the best possible lower bound that could be obtained by our approach is $s(\text{IP2}_n) \geq z_\infty^n$ where $z_\infty := \lim_{k \rightarrow \infty} (z_k)^{1/k}$. However, finding the value of z_∞ was left as an open problem in [1].

Table 2. The values of z_k and $z_k^{1/k}$ for $k \leq 10$. The numbers shown in the table are truncated (not rounded) at the third or fourth decimal places.

n	2	3	4	5	6	7	8	9	10
z_k	1.5	2	2.833	4.027	5.750	8.254	11.970	17.335	25.207
$z_k^{1/k}$	1.2247	1.2599	1.2974	1.3213	1.3384	1.3519	1.3638	1.3729	1.3808

4.2 New Lower Bounds on THR ◦ SYM Circuits

In this section, we show that $z_\infty \geq 1.5$ establishing a new lower bound on the size complexity of IP_{2^n} for THR ◦ SYM circuits.

Theorem 10. *For every $k \geq 1$,*

$$z_k \geq 1.5^k \left(1 - \frac{1}{\sqrt{k}}\right)^k.$$

Hence, $s(\text{IP}_{2^n}) = \Omega((1.5 - \epsilon)^n)$ for every $\epsilon > 0$.

Although we only prove the lower bound, we strongly believe that our bound on z_∞ is tight, i.e., $z_\infty = 1.5$. Note that $s(\text{IP}_{2^n}) \leq 2^n$ by the construction described in Introduction and the fact that AND is contained in SYM. To the best of our knowledge, this is the best known upper bound on $s(\text{IP}_{2^n})$ ¹.

Proof of Theorem 10. The proof is done by giving a feasible solution to the dual of the LP problem (3), and then estimating the value of the objective function.

We define Z_k to be

$$Z_k := \{\{2i + a, 2j + b\} \mid i, j \in [k], i \neq j \text{ and } a, b \in \{0, 1\}\}.$$

For $\mathbf{x} \in \{0, 1\}^{2k}$ and $v \in [2k]$, let x_v denote the v 's bit of $\mathbf{x}$.

The dual of (3) is given by

$$\begin{aligned} & \text{Maximize } z_{k-1} \sum_{v \in [2k]} s_v + z_{k-2} \sum_{\{u,v\} \in Z_k} t_{u,v}, \\ & \text{Subject to } \sum_{v \in [2k]: x_v = 1} s_v + \sum_{\{u,v\} \in Z_k: x_u \neq x_v} t_{u,v} \leq 1 \& (\mathbf{x} \in \{0, 1\}^{2k}), \\ & \quad s_v \geq 0 \quad (v \in [2k]), \\ & \quad t_{u,v} \geq 0 \quad (\{u, v\} \in Z_k). \end{aligned} \tag{6}$$

The LP duality theorem guarantees that the maximum value of the objective function in this dual program (6) equals to z_k . Since LP (6) is a maximization problem, any feasible solution gives a lower bound on z_k .

Here we present a feasible solution to LP (6) that will be analyzed in the proof. Define

$$\mathbf{s} \circ \mathbf{t} = (s_v)_{v \in [2k]} \circ (t_{u,v})_{\{u,v\} \in Z_k} \in \mathbb{R}^{2k+4\binom{k}{2}}$$

as follows: For $v = 1, \dots, 2k$, let $s_v = \frac{3}{4k}$ if v is odd and $s_v = 0$ if v is even. For $\{u, v\} \in Z_k$, let $t_{u,v} = \frac{9}{4k^2}$ if both of u and v are odd and $t_{u,v} = 0$ otherwise. Note that we inspired this solution through actually solving LP (6) using an LP solver.

In order to show the feasibility of $\mathbf{s} \circ \mathbf{t}$, it is enough to verify that the first constraint in LP (6) is satisfied. For $\mathbf{x} \in \{0, 1\}^{2k}$, let

$$\alpha_{\mathbf{x}} = \frac{|\{v \mid v \in \{1, 3, 5, \dots, 2k-1\} \text{ and } x_v = 1\}|}{k}.$$

¹ Actually, this is true only in an asymptotic sense. For example, an exhaustive computation shows $s(\text{IP}_{2^2}) = 2$, $s(\text{IP}_{2^3}) \leq 4$, $s(\text{IP}_{2^4}) \leq 7$ and $s(\text{IP}_{2^5}) \leq 14$.

Then, for $\mathbf{x} \in \{0, 1\}^{2k}$, the first constraint in LP (6) can be written as

$$\frac{3}{4k}\alpha_{\mathbf{x}}k + \frac{9}{4k^2}\alpha_{\mathbf{x}}k(1 - \alpha_{\mathbf{x}})k - 1 \leq 0.$$

This can easily be verified by observing that the LHS of this equation is equal to $-\left(\frac{3}{2}\alpha_{\mathbf{x}} - 1\right)^2$, completing the proof of the feasibility of $s \circ t$.

We proceed to the estimation of the value of the objective function.

The proof is by the induction on k . For $k \leq 10$, we can verify the theorem by a direct calculation (see Table 2). Suppose that $k \geq 11$. By the definition of $s \circ t$ and the inductive assumption, we have

$$\begin{aligned} z_k &\geq z_{k-1} \frac{3}{4k}k + z_{k-2} \frac{9}{4k^2} \binom{k}{2} \\ &\geq \frac{3}{4} \cdot 1.5^{k-1} \left(1 - \frac{1}{\sqrt{k-1}}\right)^{k-1} + \frac{9}{8} \cdot 1.5^{k-2} \left(1 - \frac{1}{\sqrt{k-2}}\right)^{k-2} \left(1 - \frac{1}{k}\right) \\ &= \frac{1}{2} \cdot 1.5^k \left\{ \left(1 - \frac{1}{\sqrt{k-1}}\right)^{k-1} + \left(1 - \frac{1}{\sqrt{k-2}}\right)^{k-2} \left(1 - \frac{1}{k}\right) \right\}. \end{aligned} \quad (7)$$

By an elementary but somewhat lengthy calculation, we can show that

$$z_k \geq 1.5^k \left(1 - \frac{1}{\sqrt{k}}\right)^k$$

as desired. The detailed calculations are omitted due to the page restriction and will appear in the full version of the paper. $\square$

Acknowledgement. The author would like to thank anonymous referees for their helpful comments and suggestions. This work is supported in part by JSPS Kakenhi 18K11152 and 18H04090.

References

1. Amano, K., Maruoka, A.: On the complexity of depth-2 circuits with threshold gates. In: Jędrzejowicz, J., Szepietowski, A. (eds.) MFCS 2005. LNCS, vol. 3618, pp. 107–118. Springer, Heidelberg (2005). https://doi.org/10.1007/11549345_11
2. Amano, K., Tate, S.: On XOR lemmas for the weight of polynomial threshold functions. *Inf. Comput.* **269**, 104439 (2019)
3. Basu, S., Bhatnagar, N., Gopalan, P., Lipton, R.: Polynomials that sign represent parity and descartes' rule of signs. *Comput. Complex.* **17**(3), 377–406 (2008). (Conference version in CCC '04)
4. Bruck, J.: Harmonic analysis of polynomial threshold functions. *SIAM J. Discrete Math.* **3**(2), 168–177 (1990)
5. Chattopadhyay, A., Mande, N.: A short list of equalities induces large sign rank. In: *Proceedings of FOCS 2018*, pp. 47–58 (2018)
6. Eldan, R., Shamir, O.: The power of depth for feedforward neural networks. In: *Proceedings of COLT 2016*, pp. 907–940 (2016)

7. Forster, J.: A linear lower bound on the unbounded error probabilistic communication complexity. *J. Comput. Syst. Sci.* **65**(4), 612–625 (2002)
8. Forster, J., Krause, M., Lokam, S.V., Mubarakzjanov, R., Schmitt, N., Simon, H.U.: Relations between communication complexity, linear arrangements, and computational complexity. In: Hariharan, R., Vinay, V., Mukund, M. (eds.) *FSTTCS 2001*. LNCS, vol. 2245, pp. 171–182. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45294-X_15
9. Goldmann, M., Håstad, J., Razborov, A.: Majority gates vs. general weighted threshold gates. *Comput. Complex.* **2**, 277–300 (1992)
10. Hajnal, A., Maass, W., Pudlák, P., Szegedy, M., Turán, G.: Threshold circuits of bounded depth. *J. Comput. Syst. Sci.* **46**(2), 129–154 (1993)
11. Hansen, K., Podolskii, V.: Exact threshold circuits. In: *Proceedings of CCC 2010*, pp. 270–279 (2010)
12. Jukna, S.: *Boolean Function Complexity, Advances and Frontiers. Algorithms and Combinatorics*, vol. 27. Springer, Heidelberg (2012). <https://doi.org/10.1007/978-3-642-24508-4>
13. Kane, D., Williams, R.: Super-linear gate and super-quadratic wire lower bounds for depth-two and depth-three threshold circuits. In: *Proceedings of STOC 2016*, pp. 633–643 (2016)
14. Krause, M., Pudlák, P.: On the computational power of depth-2 circuits with threshold and modulo gates. *Theor. Comput. Sci.* **174**(1–2), 137–156 (1997)
15. Muroga, S.: *Threshold Logic and Its Applications*. Wiley, Hoboken (1971)
16. Nisan, N.: The communication complexity of threshold gates. In: *Proceedings of “Combinatorics, Paul Erdős is Eighty”*, pp. 301–315 (1994)
17. Roychowdhury, V., Orlitsky, A., Siu, K.Y.: Lower bounds on threshold and related circuits via communication complexity. *IEEE Trans. Inf. Theory* **40**(2), 467–474 (1994)
18. Safran, I., Eldan, R., Shamir, O.: Depth separations in neural networks: what is actually being separated?. In: *Proceedings of COLT 2019*, pp. 2664–2666 (2019). (full version at [Arxiv:1904.06984](https://arxiv.org/abs/1904.06984))
19. Sezener, C., Oztop, E.: Minimal sign representation of boolean functions: algorithms and exact results for low dimensions. *Neural Comput.* **27**(8), 1796–1823 (2015)