Vol.2020-AL-176 No.5
2020/1/30

数理計画を用いた閾値回路の計算複雑さの解析

天野　一幸[1,a)]

概要：2 層の閾値回路を用いて GF(2) 上の内積関数 $\mathsf{IP2}_n(x_1,\ldots,x_n,y_1,\ldots,y_n) := \sum_i x_i y_i \pmod 2$ を計算する場合の回路サイズについて，以下の 2 点を証明する．(i) $\mathsf{IP2}_n$ に対する回路サイズの上界 $O(1.682^n)$，および，出力閾値素子の重みを多項式に限定した場合の回路サイズの上界 $O(1.899^n)$．(ii) 入力段側の素子を対称関数素子に限定した 2 層回路で $\mathsf{IP2}_n$ を計算する回路サイズに対する下界 $\Omega((1.5-\epsilon)^n)$．結果 (i) は，ある性質を満たす入力数が小さな回路から n 入力回路が帰納的に構成できることを示し，所望の回路を求める整数計画問題を実際に解くことにより得られる．結果 (ii) は著者が過去の研究 [MFCS '05] で提案した，回路サイズ下界を得ることのできる線形計画問題に対し，その双対問題の解を明示的に与えることにより示される．

On the size of depth-two threshold circuits for the inner product mod 2 function

Kazuyuki Amano[1,a)]

Abstract: In this report, we study the size of depth-two threshold circuits computing the inner product mod 2 function $\mathsf{IP2}_n(x_1,\ldots,x_n,y_1,\ldots,y_n) := \sum_i x_i y_i \pmod 2$. First, we reveal that $\mathsf{IP2}_n$ 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 $\mathsf{THR}\circ\mathsf{THR}$ 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 $\mathsf{MAJ}\circ\mathsf{THR}$ 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 $\mathsf{THR}\circ\mathsf{SYM}$ circuit). We show that any such circuit computing $\mathsf{IP2}_n$ 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.

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 $\mathsf{THR}\circ\mathsf{THR}$ 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 $\mathsf{THR}\circ\mathsf{THR}$ 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:

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

The inner product mod 2 function $\mathsf{IP2}_n$ 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 $\mathsf{IP2}_n$ has a

[1] 群馬大学
Tenjin 1-5-1, Kiryu, Gunma 376–8515, Japan
This article is a draft of a paper that will be appeared in LATA 2020.
[a)] amano@gunma-u.ac.jp

ⓒ 2020 Information Processing Society of Japan

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

表 1 Known upper and lower bounds on the size of depth-two circuits using threshold gates that computes $\mathsf{IP2}_n$. Entries marked with (*) are shown in this paper. Unmarked results are folklore.

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 $\mathbb{Z} \cap [-\mathrm{poly}(n), \mathrm{poly}(n)]$. Hajnal et al. [10] proved that every MAJ ∘ THR circuit computing $\mathsf{IP2}_n$ 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 THR ∘ MAJ circuit computing $\mathsf{IP2}_n$ has size $\Omega(\sqrt{2}^n/\mathrm{poly}(n))$ by lowerbounding the sign-rank of the communication matrix of $\mathsf{IP2}_n$.

Note that $\mathsf{IP2}_n$ 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 $\mathsf{IP2}_n$ 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.

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 $\mathsf{IP2}_n$. Although we know that lower bounds are more preferable, we pursuit 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 $\mathsf{IP2}_n$ can be computed by a THR ∘ AND circuit (hence also by a THR ∘ THR circuit) of size 2^n by applying the inclusion-exclusion formula. Namely,

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

To the best of our knowledge, no asymptotically better bound has not been published. Note that $\mathsf{IP2}_n$ has $2n$ input variables and the construction via the DNF representation of $\mathsf{IP2}_n$ needs $\sim 3^n$ gates.

In this work, we show that $\mathsf{IP2}_n$ has a depth-two threshold circuit of size significantly smaller than 2^n. Namely, we give an explicit construction of a THR ∘ THR circuit of size $O(1.682^n)$ as well as a MAJ ∘ THR circuit of size $O(1.899^n)$ computing $\mathsf{IP2}_n$.

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 THR ∘ 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 THR ∘ SYM circuits for $\mathsf{IP2}_n$.

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 THR ∘ SYM circuits for $\mathsf{IP2}_n$ [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 THR ∘ SYM circuits for $\mathsf{IP2}_n$. 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 Section 2, we introduce some notations. In Section 3, we give new

ⓒ 2020 Information Processing Society of Japan

upper bounds on the size of depth-two threshold circuits for $\mathsf{IP2}_n$. Then in Section 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 $\mathsf{THR} \circ \mathsf{SYM}$ circuits for $\mathsf{IP2}_n$.

2. Preliminaries

For an integer $n \geq 1$, $[n]$ denotes the set $\{1, 2, \ldots, n\}$. The *inner product mod 2 function* $\mathsf{IP2}_n$ is a Boolean function over $2n$ variables defined by

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

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

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

$$[\![w_1 x_1 + \cdots + w_n x_n \geq t]\!],$$

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

$$[\![w_1 x_1 + \cdots + w_n x_n = t]\!].$$

We call $w_1, \ldots, w_n$ 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 $\mathcal{C}$, and every bottom gate computes a function in $\mathcal{D}$ is called a $\mathcal{C} \circ \mathcal{D}$ *circuit*. For example, a $\mathsf{THR} \circ \mathsf{THR}$ 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 $\mathcal{C} \circ \mathcal{D}$ circuits is the minimum size of a $\mathcal{C} \circ \mathcal{D}$ 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

$$[\![x_1 - 2x_2 + 3x_3 \geq 2]\!]$$

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 $f : \{0, 1\}^n \to \mathbb{R}$ 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 Section 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 $\mathsf{IP2}_n$, 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 $\mathsf{THR} \circ \mathsf{ETHR}$ circuit of size s. Then, f can be computed by a $\mathsf{THR} \circ \mathsf{THR}$ circuit of size at most $2s$. The same relationship holds for $\mathsf{MAJ} \circ \mathsf{ETHR}$ and $\mathsf{MAJ} \circ \mathsf{THR}$ 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 $\mathsf{IP2}_2(x_1, x_2, y_1, y_2)$. Define two exact threshold functions g_1 and g_2 as follows.

$$g_1(x_1, x_2, y_1, y_2) = [\![x_1 + x_2 + y_1 + y_2 = 1]\!],$$
$$g_2(x_1, x_2, y_1, y_2) = [\![x_1 - x_2 + y_1 - y_2 = 0]\!].$$

It is easy to verify that

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

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

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

ⓒ 2020 Information Processing Society of Japan

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

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 THR ∘ ETHR circuit of size $O(3^{n/2}) = O(1.733^n)$ for $\mathsf{IP2}_n$, and also have a THR ∘ THR circuit of the same order by Lemma 1.

It is natural to expect that we can obtain a better bound by considering $\mathsf{IP2}_k$ 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 $\boldsymbol{x} = (x_1,\ldots,x_k) \in \{0,1\}^k$ and $\boldsymbol{y} = (y_1,\ldots,y_k) \in \{0,1\}^k$. Suppose that $\mathsf{IP2}_k$ can be represented by the sign of the linear combination of ℓ exact threshold functions where all weights are integers, i.e.,

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

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

(1) The size complexity of $\mathsf{IP2}_n$ for THR ∘ ETHR circuits as well as THR ∘ THR circuits is $O((\ell^{1/k})^n)$,

(2) The size complexity of $\mathsf{IP2}_n$ for MAJ ∘ THR circuits is $O((\sum_{i\in[\ell]}|w_i|)^{n/k})$.

Proof (sketch) First, observe that $\mathsf{IP2}_n$ is just a PARITY of n/k copies of $\mathsf{IP2}_k$. Replace each $\mathsf{IP2}_k$ with a constructed ℓ-gate THR∘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. The proof for Statement 2 is similar. □

With the aid of computers, we found a formula of length 8 for $\mathsf{IP2}_4$ as well as a formula of total weight 13 for $\mathsf{IP2}_4$ that lead us to the following theorems.

Theorem 4 The size complexity of $\mathsf{IP2}_n$ for THR ∘ ETHR circuits (and also for THR ∘ THR circuits) is $O(8^{n/4}) = O(1.682^n)$.

Theorem 5 The size complexity of $\mathsf{IP2}_n$ for MAJ ∘ THR circuits is $O(13^{n/4}) = O(1.899^n)$.

Proof of Theorem 4. Let $\{x_1,\ldots,x_4,y_1,\ldots,y_4\}$ denote the input variables for $\mathsf{IP2}_4$. For $i \in [4]$, we write $z_i := x_i + y_i$. We introduce the following seven exact threshold functions and write them as $g_1,\ldots,g_7$.

$$[\![-z_1+z_2+z_3+z_4=1]\!],\quad [\![z_1-z_2+z_3+z_4=1]\!],$$
$$[\![z_1+z_2-z_3+z_4=1]\!],\quad [\![z_1+z_2+z_3-z_4=1]\!],$$
$$[\![z_1-z_2-z_3+z_4=0]\!],\quad [\![z_1-z_2+z_3-z_4=0]\!],$$
$$[\![z_1+z_2-z_3-z_4=0]\!].$$

It is elementary to verify that

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

This gives a desired bound by Theorem 3. □

Proof of Theorem 5. Let $\{x_1,\ldots,x_4,y_1,\ldots,y_4\}$ denote the input variables for $\mathsf{IP2}_4$. For $i \in [4]$, we write $z_i := x_i + y_i$. We introduce the following twelve exact threshold functions and write them as $g_1,\ldots,g_{12}$.

$$[\![3z_1-3z_2+2z_3+4z_4=8]\!],$$
$$[\![3z_1-3z_2+4z_3+2z_4=8]\!],$$
$$[\![-3z_1+3z_2+2z_3+4z_4=8]\!],$$
$$[\![-3z_1+3z_2+4z_3+2z_4=8]\!],$$
$$[\![2z_1+4z_2+3z_3-3z_4=8]\!],$$
$$[\![4z_1+2z_2+3z_3-3z_4=8]\!],$$
$$[\![2z_1+4z_2-3z_3+3z_4=8]\!],$$
$$[\![4z_1+2z_2-3z_3+3z_4=8]\!],$$
$$[\![3z_1+3z_2+2z_3+4z_4=11]\!],$$
$$[\![3z_1+3z_2+4z_3+2z_4=11]\!],$$
$$[\![2z_1+4z_2+3z_3+3z_4=11]\!],$$
$$[\![4z_1+2z_2+3z_3+3z_4=11]\!].$$

It is elementary to verify that

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

This gives a desired bound by Theorem 3. □

It is plausible that our bounds can further be improved by considering $\mathsf{IP2}_k$ for $k \geq 5$ as a base case. We remark that, for the case of MAJ ∘ 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 MAJ ∘ ETHR circuits for $\mathsf{IP2}_n$. 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 $\mathsf{IP2}_n$ for MAJ ∘ ETHR circuits. Currently, we do not know such a barrier for THR ∘ THR circuits.

4. Lower Bounds for THR ∘ SYM Circuits

In this section, we show $\Omega((1.5-\epsilon)^n)$ lower bounds on the size of depth-two circuits for $\mathsf{IP2}_n$ where the top gate is a threshold gate and the bottom gates are symmetric gates. In Section 4.1, we review our LP-based method presented in our previous work [1], and then we establish the lower bound in Section 4.2.

ⓒ 2020 Information Processing Society of Japan

Throughout this section, we label the input variables of $\mathsf{IP2}_n$ as $\{x_1, \ldots, x_{2n}\}$ and define $\mathsf{IP2}_n(x_1, \ldots, x_{2n}) := \sum_{i\in[n]} x_{2i-1}x_{2i} \pmod 2$. 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 $\mathsf{THR} \circ \mathsf{SYM}$ circuit. For a Boolean function f, the size complexity of $\mathsf{IP2}_n$ for $\mathsf{THR} \circ \mathsf{SYM}$ circuits is simply denoted by $s(f)$. Throughout of this section, we treat a $\mathsf{THR} \circ \mathsf{SYM}$ circuit as the sign of a real polynomial.

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

$$f(x_1, \ldots, x_n) = 0 \implies P(x_1, \ldots, x_n) > 0,$$
$$f(x_1, \ldots, x_n) = 1 \implies P(x_1, \ldots, x_n) < 0. \qquad \square$$

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

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

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 $\mathrm{supp}(P) := \{S \subseteq X \mid w_S \neq 0\}$. 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 $\mathsf{IP2}_n$ is $X := \{x_1, \ldots, x_{2n}\}$.

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 $X_k = \{x_1, x_2, \ldots, x_{2k}\}$ 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.

Minimize:
$$\sum_{T\subseteq X_k} q_T,$$

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

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 $s(\mathsf{IP2}_n) \geq z_{k-1} \cdot s(\mathsf{IP2}_{n-(k-1)})$ and $s(\mathsf{IP2}_n) \geq z_{k-2} \cdot s(\mathsf{IP2}_{n-(k-2)})$ for every n. Let z_k be the minimum value of the objective function of the LP problem (3). Then $s(\mathsf{IP2}_n) \geq z_k \cdot (\mathsf{IP2}_{n-k})$

The following corollary is immediate from Fact 7.

Corollary 8 ([1]) For every $k \geq 1$, $s(\mathsf{IP2}_n) \geq (z_k^{1/k})^n$. $\square$

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

Let $f : \{0,1\}^X \to \mathbb{R}$ be a real function and $\rho : X \to \{0,1,*\}$ be a partial assignment to X. Let $\mathrm{res}(\rho)$ denote the set of variables that assigned a constant by ρ, i.e., $\mathrm{res}(\rho) := \{v \in X \mid \rho(v) \neq *\}$. The restriction of f by ρ, denoted by $f|_\rho$, is the function obtained by setting x_i to $\rho(x_i)$ if $x_i \in \mathrm{res}(\rho)$ 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 $S - \mathrm{res}(\rho)$. Then, if there are two (or more) functions $h_{S_1}|_\rho$ and $h_{S_2}|_\rho$ such that $S_1\backslash\mathrm{res}(\rho) = S_2\backslash\mathrm{res}(\rho)$, 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

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

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

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

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 $\mathrm{res}(\rho_1) = \mathrm{res}(\rho_2)$. Then, $\sum_{v\in T\cap\mathrm{res}(\rho_1)} \rho_1(v) = \sum_{v\in T\cap\mathrm{res}(\rho_2)} \rho_2(v)$ 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 $\mathsf{IP2}_n$. 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 $\mathrm{res}(\rho_1) = \mathrm{res}(\rho_2) = \{x_{2i-1}, x_{2i}\}$, $(\rho_1(v), \rho_1(u)) = (0,1)$ 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 $\mathsf{IP2}_n|_{\rho_1} \equiv \mathsf{IP2}_{n-1}$ and $\mathsf{IP2}_n|_{\rho_2} \equiv \overline{\mathsf{IP2}_{n-1}}$. This implies that the polynomial $P|_{\rho_1} - P|_{\rho_2}$ sign represents $\mathsf{IP2}_{n-1}$. Fact 9 says that $P_T|_{\rho_1} - P_T|_{\rho_2}$ is vanished if $v \notin T$. Hence, we have

$$\begin{aligned}\sum_{T:v\in T} \tilde{q}_T &\geq |\mathrm{supp}(P|_{\rho_1} - P|_{\rho_2})| \\ &\geq s(\mathsf{IP2}_{n-1}) \\ &\geq z_{k-1} \cdot s(\mathsf{IP2}_{n-k}), \end{aligned} \tag{4}$$

where the last inequality follows from the assumption in the statement of Fact 7. Let $q_T := \tilde{q}_T / s(\mathsf{IP2}_{n-k})$ for $T \subseteq X_k$. By dividing both side of (4) by $s(\mathsf{IP2}_{n-k})$, 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 $\mathrm{res}(\rho_1) = \mathrm{res}(\rho_2) = \{x_{2i-1}, x_{2i}, x_{2j-1}, x_{2j}\}$, $(\rho_1(v), \rho_1(v'), \rho_1(u), \rho_1(u')) = (0, 1, 1, 0)$ 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 $\mathsf{IP2}_n|_{\rho_1} \equiv \mathsf{IP2}_{n-2}$ and $\mathsf{IP2}_n|_{\rho_2} \equiv \overline{\mathsf{IP2}_{n-2}}$, and hence $P|_{\rho_1} - P|_{\rho_2}$ sign represents $\mathsf{IP2}_{n-2}$. In addition, $P_T|_{\rho_1} - P_T|_{\rho_2}$ is vanished if $|T \cap \{u, v\}|$ is zero or two. Hence, we have

$$\begin{aligned}\sum_{T:|\{u,v\}\cap T|=1} \tilde{q}_T &\geq |\mathrm{supp}(P|_{\rho_1} - P|_{\rho_2})| \\ &\geq s(\mathsf{IP2}_{n-2}) \\ &\geq z_{k-2} \cdot s(\mathsf{IP2}_{n-k}), \end{aligned} \tag{5}$$

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

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

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

If P is an optimal polynomial for $\mathsf{IP2}_n$, then

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

which is equivalent to

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

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

n	2	3	4	5	6	7
z_k	1.5	2	2.833	4.027	5.750	8.254
$z_k^{1/k}$	1.2247	1.2599	1.2974	1.3213	1.3384	1.3519

$\cdots$	8	9	10
	11.970	17.335	25.207
	1.3638	1.3729	1.3808

表 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.

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 $s(\mathsf{IP2}_n) = \Omega(\sqrt{2}^n/n)$ due to Forster et al. [7], [8].

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

4.2 New Lower Bounds on $\mathsf{THR} \circ \mathsf{SYM}$ Circuits

In this section, we show that $z_\infty \geq 1.5$ establishing a new lower bound on the size complexity of $\mathsf{IP2}_n$ for $\mathsf{THR} \circ \mathsf{SYM}$ circuits.

Theorem 10 For every $k \geq 1$,

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

Hence, $s(\mathsf{IP2}_n) = \Omega((1.5 - \epsilon)^n)$ 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 $s(\mathsf{IP2}_n) \leq 2^n$ 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 $s(\mathsf{IP2}_n)$[*1].

Proof of Theorem 10 (sketch). 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 $\boldsymbol{x} \in \{0, 1\}^{2k}$ and $v \in [2k]$, let x_v denote the v's bit of $\boldsymbol{x}$.

The dual of (3) is given by

[*1] Actually, this is true only in an asymptotic sense. For example, an exhaustive computation shows $s(\mathsf{IP2}_2) = 2$, $s(\mathsf{IP2}_3) \leq 4$, $s(\mathsf{IP2}_4) \leq 7$ and $s(\mathsf{IP2}_5) \leq 14$.

ⓒ 2020 Information Processing Society of Japan

Maximize:

$$z_{k-1}\sum_{v\in[2k]} s_v + z_{k-2}\sum_{\{u,v\}\in Z_k} t_{u,v},$$

Subject to: (6)

$$\begin{aligned}
\sum_{v\in[2k]:x_v=1} s_v + \sum_{\{u,v\}\in Z_k:x_u\neq x_v} t_{u,v} &\leq 1 \quad (\boldsymbol{x}\in\{0,1\}^{2k}),\\
s_v &\geq 0 \quad (v\in[2k]),\\
t_{u,v} &\geq 0 \quad (\{u,v\}\in Z_k).
\end{aligned}$$

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

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

as follows: For $v=1,\ldots,2k$, 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 $\boldsymbol{s}\circ\boldsymbol{t}$, it is enough to verify that the first constraint in LP (6) is satisfied. For $\boldsymbol{x}\in\{0,1\}^{2k}$, let

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

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

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

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

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 $\boldsymbol{s}\circ\boldsymbol{t}$ and the inductive assumption, we have

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

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 in this version.

参考文献

[1] K. Amano and A. Maruoka. On the complexity of depth-2 circuits with threshold gates. *Proc. of MFCS '05, LNCS*, 3618:107 – 118, 2005.

[2] K. Amano and S. Tate. On XOR lemmas for the weight of polynomial threshold functions. *Inf. Comput.*, 269:104439, 2019.

[3] S. Basu, N. Bhatnagar, P. Gopalan, and R.J. Lipton. Polynomials that sign represent parity and descartes' rule of signs. *Computational Complexity*, 17(3):377–406, 2008. (Conference version in CCC '04).

[4] J. Bruck. Harmonic analysis of polynomial threshold functions. *SIAM J. Discrete Math.*, 3(2):168–177, 1990.

[5] A. Chattopadhyay and N.S. Mande. A short list of equalities induces large sign rank. *Proc. of FOCS '18*, pages 47–58, 2018.

[6] R. Eldan and O. Shamir. The power of depth for feedforward neural networks. *Proc. of COLT '16*, pages 907–940, 2016.

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

[8] J. Forster, M. Krause, S.V. Lokam, R. Mubarakzjanov, N. Schmitt, and H.U. Simon. Relations between communication complexity, linear arrangements and computational complexity. *Prof. of FSTTCS '01*, pages 171–182, 2001.

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

[10] 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.

[11] K.A. Hansen and V.V. Podolskii. Exact threshold circuits. *Proc. of CCC '10*, pages 270–279, 2010.

[12] S. Jukna. *Boolean function complexity - Advances and frontiers*, volume 27 of *algorithms and combinatorics*. Springer, 2012.

[13] D.M. Kane and R. Williams. Super-linear gate and super-quadratic wire lower bounds for depth-two and depth-three threshold circuits. *Proc. of STOC '16*, pages 633–643, 2016.

[14] 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.

[15] S. Muroga. *Threshold logic and its applications*. John Wiley & Sons, Inc., 1971.

[16] N. Nisan. The communication complexity of threshold gates. *Proc. of "Combinatorics, Paul Erdős is Eighty"*, pages 301–315, 1994.

[17] V.P. Roychowdhury, A. Orlitsky, and K-Y. Siu. Lower bounds on threshold and related circuits via communication complexity. *IEEE Trans. on Inf. Theory*, 40(2):467–474, 1994.

[18] I. Safran, R. Eldan, and O. Shamir. Depth separations in neural networks: what is actually being separated? *Proc. of COLT '19*, pages 2664–2666, 2019. (full version at `Arxiv:1904.06984`).

[19] C.E. Sezener and E. Oztop. Minimal sign representation of boolean functions: Algorithms and exact results for low dimensions. *Neural Computation*, 27(8):1796–1823, 2015.

ⓒ 2020 Information Processing Society of Japan