

Energy-Efficient Threshold Circuits Computing Mod Functions

Akira Suzuki Kei Uchizawa* Xiao Zhou

Graduate School of Information Sciences, Tohoku University
Aramaki-aza Aoba 6-6-05, Aoba-ku, Sendai, 980-8579, Japan.
{a.suzuki, uchizawa, zhou}@ecei.tohoku.ac.jp

Abstract

We prove that the modulus function MOD_m of n variables can be computed by a threshold circuit C of energy e and size $s = O(e(n/m)^{1/(e-1)})$ for any integer $e \geq 2$, where the energy e is defined to be the maximum number of gates outputting “1” over all inputs to C , and the size s to be the number of gates in C . Our upper bound on the size s almost matches the known lower bound $s = \Omega(e(n/m)^{1/e})$.

1 Introduction

Neuronal signals play fundamental role in information processing of the brain. A neuron emitting a signal is said to be “firing.” Recent biological studies report the following fact about the energy consumption of the neuronal firing: the energy cost of a neuronal firing is high while energy supplied to the brain is limited, and hence neural networks must have low firing activity [1, 4]. Consequently, many neuroscientists consider that the metabolic limit must influence the way in which information is processed, and the brain have countered this metabolic constraint by adopting energy-efficient circuit designs [2, 3, 6, 14]. Uchizawa, Douglas and Maass consider the problem posed above from the view point of theoretical computer science, and introduce a new complexity measure called the energy complexity of threshold circuits [10], where a threshold circuit, which is a combinatorial circuit consisting of threshold gates, is a theoretical model of neural circuit [5, 7, 8, 9]. Based on the biological fact above, the *energy* e of a threshold circuit C is defined as the maximum number of threshold gates outputting “1” over all inputs to C . In previous research, several facts are known on the computational power of threshold circuits with small energy [10, 11, 12, 13]. Particularly in the paper [10], they find a non-trivial circuit structure that benefit energy-efficiency, and provide threshold circuits of polynomial size and energy $O(\log n)$ for a fairly large class of Boolean functions of n variables. However, their construction is not specialized for a particular task, and hence it sometimes gives a redundant circuit.

In this paper, we consider one of the fundamental and well-studied Boolean functions in the theory of circuit complexity, the modulus function, as a particular

*Supported by MEXT Grant-in-Aid for Young Scientists (B) No.21700003

task, where the modulus function MOD_m of n variables for two positive integers m and n is defined as follows: $\text{MOD}_m(\mathbf{x}) = 0$ if the number of “1”s in an input $\mathbf{x} \in \{0, 1\}^n$ is a multiple of m and, otherwise, $\text{MOD}_m(\mathbf{x}) = 1$. Although the modulus function may be far from real tasks that neural networks in the brain perform, we believe that considering such a simple and fundamental task makes an important step for understanding what circuit structure benefits the energy-efficiency of threshold circuits. In [13], it is proved that size and energy of a threshold circuit computing the modulus function cannot be simultaneously small: Any threshold circuit C of energy e computing MOD_m of n variables has size

$$s = \Omega \left(e \left(\frac{n}{m} \right)^{1/e} \right). \quad (1)$$

We prove in this paper that MOD_m of n variables can be computed by a threshold circuit of energy e and size

$$s = O \left(e \left(\frac{n}{m} \right)^{1/(e-1)} \right) \quad (2)$$

for every integer $e \geq 2$. Comparing the right-hand side of Eq. (1) with one of Eq. (2), we can find the difference between the terms only in the exponent of n/m . Thus, our upper bound almost matches the lower bound, and implies that there exists a tight tradeoff between size and energy of threshold circuits computing modulus function. We obtain the result by construction of the desired threshold circuits, and hence it exhibits a circuit design of energy-efficient threshold circuits.

In addition, we consider an extreme case where threshold circuits have energy $e = 1$. In this case, we prove that any threshold circuit C computing the PARITY of n variables must have an exponential number of gates in n . On the other hand, Eq. (2) implies that PARITY (i.e., MOD_2) can be computed by a threshold circuit of size $s = O(n)$ and energy $e = 2$. Thus, we know from these facts that there exists a significant gap of computational power between threshold circuits of $e = 1$ and ones of $e = 2$.

The rest of the paper is organized as follows. In Section 2, we define some terms on threshold circuits and the modulus function. In Section 3, we first provide our main theorem. We then give a technical lemma, and prove the theorem using the lemma. In Section 4, we prove the technical lemma given in Section 3. In Section 5, we give the lower bound for threshold circuits of energy one.

2 Preliminaries

A *threshold circuit* C is a combinatorial circuit of threshold gates, and is expressed by a directed acyclic graph. Let n be the number of input variables to C . Then each node of in-degree 0 in C corresponds to one of the n input variables x_1, x_2, \dots, x_n , and the other nodes correspond to threshold gates. We define the *size* $s(C)$ of a threshold circuit C as the number of threshold gates in C .

Let $g_1^C, g_2^C, \dots, g_{s(C)}^C$ be the gates in a threshold circuit C . For each gate g_i^C , $1 \leq i \leq s(C)$, and let $\mathbf{z}(\mathbf{x}) = (z_1(\mathbf{x}), z_2(\mathbf{x}), \dots, z_{k_i}(\mathbf{x})) \in \{0, 1\}^{k_i}$ be the k_i inputs of g_i^C with weights w_1, w_2, \dots, w_{k_i} and a threshold t_i for $\mathbf{x} \in \{0, 1\}^n$. Then the output $g_i^C(\mathbf{z}(\mathbf{x}))$ of the gate g_i^C is defined as follows:

$$g_i^C(\mathbf{z}(\mathbf{x})) = \text{sign} \left(\sum_{j=1}^{k_i} w_j z_j(\mathbf{x}) - t_i \right),$$

where $\text{sign}(z) = 1$ if $z \geq 0$ and $\text{sign}(z) = 0$ if $z < 0$. Simply, $g_i^C(\mathbf{z}(\mathbf{x}))$ is denoted by $g_i^C[\mathbf{x}]$. Let $f : \{0, 1\}^n \rightarrow \{0, 1\}$ be a Boolean function of n variables. Let g_s^C be a gate of out-degree 0 in C , and let the output $g_s^C[\mathbf{x}]$ of g_s be the *output* $C(\mathbf{x})$ of C . Thus, $C(\mathbf{x}) = g_s[\mathbf{x}]$ for every input $\mathbf{x} \in \{0, 1\}^n$. The gate g_s^C is called the *top gate* of C . A threshold circuit C *computes* a Boolean function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ if $C(\mathbf{x}) = f(\mathbf{x})$ for every input $\mathbf{x} \in \{0, 1\}^n$.

We define the *energy* $e(C)$ of a threshold circuit C as

$$e(C) = \max_{\mathbf{x} \in \{0, 1\}^n} \sum_{i=1}^{s(C)} g_i^C[\mathbf{x}].$$

Thus, the energy $e(C)$ is the maximum number of gates outputting “1” over all inputs $\mathbf{x} \in \{0, 1\}^n$ to C . Trivially, we have $0 \leq e(C) \leq s(C)$.

For an input variable $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \{0, 1\}^n$, we define $|\mathbf{x}|$ as the hamming weight of the inputs \mathbf{x} , that is,

$$|\mathbf{x}| = \sum_{i=1}^n x_i.$$

Then, for an integer $m \geq 2$, the modulus function MOD_m of n variables is defined as follows: For every input variable $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \{0, 1\}^n$,

$$\text{MOD}_m(\mathbf{x}) = \begin{cases} 1 & \text{if } |\mathbf{x}| \text{ is not a multiple of } m; \\ 0 & \text{otherwise.} \end{cases}$$

If $n \leq m - 1$, then $\text{MOD}_m(\mathbf{x}) = 1$ for all inputs \mathbf{x} except $\mathbf{x} = (0, 0, \dots, 0)$. Thus, a circuit consisting of a single threshold gate with weight ones for all the inputs and threshold one computes MOD_m . One may thus assume that $n \geq m$ in the remainder of the paper.

3 Energy-Efficient Circuits

Our main result is the following theorem that yields an energy-efficient threshold circuit computing the modulus function.

Theorem 1. *Let m, n be any two positive integers, and let e be any integer $e \geq 2$. Then there is a threshold circuit computing MOD_m of n variables such that its energy is at most e and its size is at most*

$$(e - 1) \left\lceil \left(\frac{n + 1}{m} \right)^{1/(e-1)} \right\rceil = O \left(e \left(\frac{n}{m} \right)^{1/(e-1)} \right). \quad (3)$$

In the paper [13], it is known that the size s and energy e of a threshold circuit C computing MOD_m of n variables cannot be simultaneously small, as described in the following theorem.

Theorem 2 ([13]). *Let C be a threshold circuit computing the function MOD_m of n variables. Then the size s and energy e of C satisfy*

$$\frac{n}{m - 1} + 1 \leq \frac{1}{\sqrt{2\pi e}} \cdot \left(\frac{2c_{npr} \cdot s}{e} \right)^e. \quad (4)$$

where $c_{npr} \cong 2.718$ is the Napier’s (or mathematical) constant.

By a simple modification of Eq. (4), we can obtain from Theorem 2 the following lower bound on the size of threshold circuits computing MOD_m of n variables.

Corollary 1. *Let C be any threshold circuit of energy e computing MOD_m of n variables. Then*

$$s(C) = \Omega\left(e\left(\frac{n}{m}\right)^{1/e}\right). \quad (5)$$

Observe the asymptotic terms in the right-hand side of Eqs. (3) and (5). We can find the difference between the terms only in the exponent of n/m : the term in Eq. (3) has $1/(e-1)$, while the term in Eq. (5) has $1/e$. Hence, the upper bound in Theorem 1 almost matches the lower bound in Corollary 1.

In the rest of the section, we prove Theorem 1. We say that a threshold circuit C is *regular* if the inputs of every gate in C includes all the inputs x_1, x_2, \dots, x_n with weight ones. In other words, every gate in C receives all the unweighted inputs x_1, x_2, \dots, x_n . The following technical lemma plays key role in our proof.

Lemma 1. *Let m, n, n' be positive integers such that $n \geq n' + 1$. Let C' be a regular threshold circuit computing MOD_m of n variables. Then, there is a regular threshold circuit C computing MOD_m of n variables such that $e(C) \leq e(C') + 1$ and*

$$s(C) \leq s(C') + \left\lceil \frac{n+1}{\lfloor \frac{n'+1}{m} \rfloor m} \right\rceil - 1.$$

We will prove the lemma in the next section. Using the lemma, we prove Theorem 1 below.

Proof of Theorem 1. Let e be an arbitrary integer at least 2. We prove the theorem by constructing a regular threshold circuit C computing MOD_m of n variables such that $e(C) \leq e$ and

$$s(C) \leq (e-1) \left\lceil \left(\frac{n+1}{m}\right)^{1/(e-1)} \right\rceil. \quad (6)$$

We provide our construction by induction on $e \geq 2$. That is, we construct a threshold circuit of energy $e+1$ from a threshold circuit of energy e . We start from the case of $e=2$ as the basis.

Basis: $e=2$.

Consider a regular threshold circuit C' consisting of a single threshold gate g with threshold $t=1$ and $m-1$ input variables. Clearly, $e(C')=1$, $s(C')=1$, and C' computes MOD_m of $n'=m-1$ variables. Therefore, Lemma 1 implies that there is a regular threshold circuit C computing MOD_m of n variables such that $e(C) \leq e(C') + 1$ and

$$s(C) \leq s(C') + \left(\left\lceil \frac{n+1}{m} \right\rceil - 1 \right).$$

Since $e(C')=1$, we have $e(C) \leq e(C') + 1 = 2 = e$. Since $s(C')=1$, we have

$$s(C) \leq s(C') + \left(\left\lceil \frac{n+1}{m} \right\rceil - 1 \right) = \left\lceil \frac{n+1}{m} \right\rceil = (e-1) \left\lceil \left(\frac{n+1}{m}\right)^{1/(e-1)} \right\rceil.$$

Inductive Step: $e \geq 3$.

By the induction hypothesis, there is a regular threshold circuit C' computing MOD_m of $n' = m\gamma^{e-1} - 1$ variables such that $e(C') \leq e$ and

$$\begin{aligned} s(C') &\leq (e-1) \left\lceil \left(\frac{n'+1}{m} \right)^{1/(e-1)} \right\rceil \\ &\leq (e-1) \left\lceil \left(\frac{(m\gamma^{e-1}-1)+1}{m} \right)^{1/(e-1)} \right\rceil \\ &\leq (e-1)\lceil \gamma \rceil \end{aligned} \tag{7}$$

for each positive integer γ . We will construct a regular threshold circuit C computing MOD_m of n variables, and show that C has the energy

$$e(C) \leq e+1 \tag{8}$$

and the size

$$s(C) \leq e \left\lceil \left(\frac{n+1}{m} \right)^{1/e} \right\rceil. \tag{9}$$

Since C' computes MOD_m of $n' = m\gamma^{e-1} - 1$ variables, by Lemma 1 there is a regular threshold circuit C computing MOD_m of n variables such that

$$e(C) \leq e(C') + 1 = e+1$$

and

$$s(C) \leq s(C') + \left\lceil \frac{n+1}{m\gamma^{e-1}} \right\rceil - 1. \tag{10}$$

We choose

$$\gamma = \left\lceil \left(\frac{n+1}{m} \right)^{1/e} \right\rceil,$$

then $\gamma^e \geq (n+1)/m$, and hence

$$\left\lceil \frac{n+1}{m\gamma^{e-1}} \right\rceil - 1 \leq \gamma. \tag{11}$$

Therefore, by Eqs. (7), (10) and (11), we have

$$\begin{aligned} s(C) &\leq s(C') + \left\lceil \frac{n+1}{m\gamma^{e-1}} \right\rceil - 1 \\ &\leq (e-1)\lceil \gamma \rceil + \gamma \\ &\leq e\lceil \gamma \rceil \\ &= e \left\lceil \left(\frac{n+1}{m} \right)^{1/e} \right\rceil. \end{aligned}$$

□

4 Proof of Lemma 1

In the section, we prove Lemma 1.

Let m, n, n' be positive integers such that $n \geq n' + 1$. Let C' be a regular threshold circuit computing MOD_m of n' variables, and $s = s(C')$. We denote by g'_1, g'_2, \dots, g'_s the threshold gates in C' . One may assume without loss of generality that g'_1, g'_2, \dots, g'_s are topologically ordered with respect to the underlying directed acyclic graph of C' , and that each gate g'_i , $1 \leq i \leq s$, receives exactly $(i-1) + n'$ inputs from the outputs of the gates $g'_1, g'_2, \dots, g'_{i-1}$ and the n' inputs $x_1, x_2, \dots, x_{n'}$. If there is some gate g'_i , $1 \leq j \leq i-1$, such that g'_i has no input from the output of g'_j , then one connects input of g'_i with weight 0 for the output of g'_j . Therefore, for each index i , $1 \leq i \leq s$, let $w'_{i,1}, w'_{i,2}, \dots, w'_{i,i-1}$ be the weights of the gate g'_i for the outputs of the gates $g'_1, g'_2, \dots, g'_{i-1}$, respectively, and denote by t'_i the threshold of g'_i . Since C' is regular, each of the gates g'_1, g'_2, \dots, g'_s has weight ones for the n' input variables. Thus, the output of the gate g'_i for each input $\mathbf{x}' \in \{0, 1\}^{n'}$ can be recursively computed by the following threshold function:

$$g'_i[\mathbf{x}'] = \begin{cases} \text{sign}(|\mathbf{x}'| - t_1) & \text{if } i = 1; \\ \text{sign}\left(|\mathbf{x}'| + \sum_{j=1}^{i-1} w'_{i,j} g'_j[\mathbf{x}'] - t'_i\right) & \text{otherwise.} \end{cases} \quad (12)$$

We show that, for any positive integer $n \geq n' + 1$, MOD_m of n variables can be computed by a regular threshold circuit C of energy $e(C) \leq e(C') + 1$ and size $s(C) \leq s(C') + \beta$, where

$$\beta = \left\lceil \frac{n+1}{\alpha m} \right\rceil - 1$$

and

$$\alpha = \left\lfloor \frac{n'+1}{m} \right\rfloor.$$

We construct the desired threshold circuit C from C' , as described below.

To obtain C , we add new input variables $x_{n'+1}, x_{n'+2}, \dots, x_n$ to C' , and connect each of the new input variables to each of the gates g'_1, g'_2, \dots, g'_s with weight one. Besides, for each index i , $1 \leq i \leq \beta$, we add a new threshold gate \hat{g}_i with weight ones for the inputs x_1, x_2, \dots, x_n and a threshold $\alpha m i$ to C' , and connect the output of the gate \hat{g}_i to the gates g'_1, g'_2, \dots, g'_s with weight $-\alpha m i$. For each index i , $1 \leq i \leq s$, we denote by g_i the gate in C that corresponds to the gate g'_i in C' , and denote by $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \{0, 1\}^n$ an input to C . Then, the output of the gate g_i for an input $\mathbf{x} \in \{0, 1\}^n$ is now represented as

$$g_i[\mathbf{x}] = \begin{cases} \text{sign}(|\mathbf{x}| - t'_1) & \text{if } i = 1; \\ \text{sign}\left(|\mathbf{x}| + \sum_{j=1}^{i-1} w'_{i,j} g_j[\mathbf{x}] - \sum_{j=1}^{\beta} \alpha m j \cdot \hat{g}_j[\mathbf{x}] - t'_i\right) & \text{otherwise.} \end{cases} \quad (13)$$

Moreover, for each index i , $2 \leq i \leq \beta$, we connect the output of the gate \hat{g}_i to the gates $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_{i-1}$ with weight $-\alpha m i$. Thus, we have

$$\hat{g}_i[\mathbf{x}] = \begin{cases} \text{sign}\left(|\mathbf{x}| - \sum_{j=i+1}^{\beta} \alpha m j \cdot \hat{g}_j[\mathbf{x}] - \alpha m i\right) & \text{if } 1 \leq i \leq \beta - 1; \\ \text{sign}(|\mathbf{x}| - \alpha m \beta) & \text{if } i = \beta \end{cases} \quad (14)$$

for $\mathbf{x} \in \{0, 1\}^n$. Clearly, C is a regular circuit, and

$$s(C) \leq s(C') + \beta = s(C') + \left\lceil \frac{n+1}{\lfloor \frac{n'+1}{m} \rfloor m} \right\rceil - 1.$$

Below we prove that C computes MOD_m of n variables, and $e(C) \leq e(C') + 1$.

Let $\mathbf{x} \in \{0, 1\}^n$ be an arbitrary input to C . Note that $0 \leq |\mathbf{x}|$ and

$$\begin{aligned} |\mathbf{x}| &\leq n \\ &\leq \alpha m \cdot \left\lceil \frac{n+1}{\alpha m} \right\rceil - 1 \\ &\leq \alpha m(\beta + 1) - 1 \end{aligned}$$

for any input $\mathbf{x} \in \{0, 1\}^n$. Let

$$i^* = \left\lfloor \frac{|\mathbf{x}|}{\alpha m} \right\rfloor,$$

then trivially

$$\alpha m i^* \leq |\mathbf{x}| \leq \alpha m(i^* + 1) - 1. \quad (15)$$

We prove the following claim.

Claim 1. *The following (i), (ii) and (iii) hold.*

- (i) $\hat{g}_i[\mathbf{x}] = 0$ for each i , $i^* + 1 \leq i \leq \beta$;
- (ii) $\hat{g}_i[\mathbf{x}] = 1$ if $i = i^*$;
- (iii) $\hat{g}_i[\mathbf{x}] = 0$ for each i , $1 \leq i \leq i^* - 1$.

In other words, if $0 \leq |\mathbf{x}| \leq \alpha m - 1$, none of $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_\beta$ outputs one; otherwise, only the gate \hat{g}_{i^*} outputs one.

Proof of Claim. For each index i , $1 \leq i \leq \beta$, let

$$p_i(\mathbf{x}) = \begin{cases} |\mathbf{x}| - \sum_{j=i+1}^{\beta} \alpha m j \cdot \hat{g}_j[\mathbf{x}] - \alpha m i & \text{if } 1 \leq i \leq \beta - 1, \\ |\mathbf{x}| - \alpha m \beta & \text{if } i = \beta. \end{cases} \quad (16)$$

Clearly, $p_i[\mathbf{x}]$ is the value in the sign function of the right hand side of Eq. (14) for $\mathbf{x} \in \{0, 1\}^n$, that is, $\hat{g}_i[\mathbf{x}] = \text{sign}(p_i(\mathbf{x}))$. We evaluate $p_i(\mathbf{x})$, and prove the following (i), (ii) and (iii).

- (i) $\hat{g}_i[\mathbf{x}] = 0$ for each i , $i^* + 1 \leq i \leq \beta$.
If $i \leq \beta - 1$, then by Eqs. (15) and (16)

$$\begin{aligned} p_i(\mathbf{x}) &\leq \alpha m(i^* + 1) - 1 - \sum_{j=i+1}^{\beta} \alpha m j \cdot \hat{g}_j[\mathbf{x}] - \alpha m i \\ &\leq \alpha m(i^* + 1 - i) - 1 \\ &\leq -1. \end{aligned} \quad (17)$$

If $i = \beta$, then by Eqs. (15) and (16) we similarly have

$$\begin{aligned} p_i(\mathbf{x}) &= |\mathbf{x}| - \alpha m \beta \\ &\leq \alpha m(i^* + 1) - 1 - \alpha m \beta \\ &\leq -1. \end{aligned} \tag{18}$$

Since $i^* + 1 \leq \beta$, Eqs. (14), (17) and (18) imply that $\hat{g}_i[\mathbf{x}] = \text{sign}(p_i(\mathbf{x})) = 0$.

(ii) $\hat{g}_i[\mathbf{x}] = 1$ if $i = i^*$.

In this case, we have $1 \leq i^* \leq \beta$. By (i) above, if $i^* \leq \beta - 1$,

$$\sum_{j=i^*+1}^{\beta} \alpha m j \cdot \hat{g}_j[\mathbf{x}] = 0,$$

and hence we have by Eqs. (15) and (16)

$$\begin{aligned} p_{i^*}(\mathbf{x}) &= |\mathbf{x}| - \alpha m i^* \\ &\geq \alpha m i^* - \alpha m i^* \\ &= 0. \end{aligned}$$

Thus Eq. (14) implies that $\hat{g}_{i^*}[\mathbf{x}] = \text{sign}(p_{i^*}(\mathbf{x})) = 1$.

(iii) $\hat{g}_i[\mathbf{x}] = 0$ for each i , $1 \leq i \leq i^* - 1$.

In this case, we have $2 \leq i + 1 \leq i^* \leq \beta$, and hence

$$\sum_{j=i+1}^{\beta} \hat{g}_j[\mathbf{x}] \geq \hat{g}_{i^*}[\mathbf{x}].$$

By (ii) above, we have $\hat{g}_{i^*}[\mathbf{x}] = 1$, and hence

$$\begin{aligned} - \sum_{j=i+1}^{\beta} \alpha m j \cdot \hat{g}_j[\mathbf{x}] &\leq -\alpha m i^* \cdot \hat{g}_{i^*}[\mathbf{x}] \\ &= -\alpha m i^*. \end{aligned} \tag{19}$$

Since $i + 1 \leq \beta$, we have $i \leq \beta - 1$. Therefore, by Eqs. (16) and (19)

$$p_i(\mathbf{x}) \leq |\mathbf{x}| - \alpha m i^* - \alpha m i. \tag{20}$$

By Eqs. (15) and (20), we have

$$\begin{aligned} p_i(\mathbf{x}) &\leq \alpha m(i^* + 1) - 1 - \alpha m i^* - \alpha m i \\ &\leq \alpha m - 1 - \alpha m i \\ &\leq -1 \end{aligned}$$

and hence Eq. (14) implies that $\hat{g}_i[\mathbf{x}] = \text{sign}(p_i(\mathbf{x})) = 0$. \square

We are now ready to prove the lemma by the claim. There are the following two cases to consider.

Case 1: $0 \leq |\mathbf{x}| \leq \alpha m - 1$.

In this case, the claim implies that none of $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_\beta$ output one. Besides, we have

$$\alpha m - 1 = \left\lfloor \frac{n' + 1}{m} \right\rfloor - 1 \leq n'.$$

Therefore, Eqs. (12) and (13) imply that, for every index i , $1 \leq i \leq s$, the output of g_i for $\mathbf{x} \in \{0, 1\}^n$ equals to the output of g'_i for an input $\mathbf{x}' \in \{0, 1\}^{n'}$ such that $|\mathbf{x}'| = |\mathbf{x}|$. Thus the number of gates outputting one is at most e . Since C' computes MOD_m , $C(\mathbf{x})$ equals to $\text{MOD}_m(\mathbf{x})$.

Case 2: $\alpha m \leq |\mathbf{x}| \leq \alpha m(\beta + 1) - 1$.

In this case, the claim implies that only the gate \hat{g}_{i^*} of $\hat{g}_1, \hat{g}_2, \dots, \hat{g}_\beta$ outputs one, and hence Eq. (13) implies that, for every index i , $1 \leq i \leq s$, the output of g_i can be represented as

$$g_i[\mathbf{x}] = \text{sign} \left(|\mathbf{x}| + \sum_{j=1}^{i-1} w'_{i,j} g_j[\mathbf{x}] - \alpha m i^* - t'_i \right). \quad (21)$$

Eq. (15) implies that

$$0 \leq |\mathbf{x}| - \alpha m i^* \leq \alpha m - 1,$$

and hence, for every index i , $1 \leq i \leq s$, we have that $g_i[\mathbf{x}]$ for $\mathbf{x} \in \{0, 1\}^n$ equals to the output of g'_i for an input $\mathbf{x}' \in \{0, 1\}^{n'}$ such that $|\mathbf{x}| - \alpha m i^* = |\mathbf{x}'|$. Thus, at most e gates of the gates g_1, g_2, \dots, g_s output one, and consequently the number of gates outputting one in C is at most $e + 1$. The circuit C' computes MOD_m of n' variables, and $|\mathbf{x}| - \alpha m j$ is a multiple of m if and only if $|\mathbf{x}'|$ is a multiple of m . Hence, $C(\mathbf{x})$ equals to $\text{MOD}_m(\mathbf{x})$.

5 Circuits of Energy One

In this section, we consider an extreme case where threshold circuits have energy $e = 1$. While we know from Theorem 1 that PARITY (i.e., MOD_2) of n variables can be computed by a threshold circuit of size $s = O(n)$ and energy $e = 2$, we can prove that any threshold circuit of energy $e = 1$ computing PARITY of n variables must have an exponential number of gates in n , as follows.

Theorem 3. *If a threshold circuit C of energy one computes PARITY of n variables, then the size s of C is at least 2^{n-1} .*

The proof is omitted due to the page limitation, but described in the appendix. Theorems 1 and 3 imply that there exists a significant gap of computational power between threshold circuits of $e = 1$ and ones of $e = 2$.

6 Conclusions

In the paper, we design energy-efficient threshold circuits computing the modulus function MOD_m , and show that MOD_m of n variables can be computed by a threshold circuit of size $s = O(e(n/m)^{1/(e-1)})$ and energy e for any integer $e \geq 2$. The upper bound on the size $s = O(e(n/m)^{1/(e-1)})$ almost matches the known lower bound $\Omega(e(n/m)^{1/e})$ presented in [13]. A Boolean function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is called symmetric if $f(\mathbf{x})$ depends only on the number of 1s in \mathbf{x} for every input $\mathbf{x} \in \{0, 1\}^n$. Thus, the modulus function is symmetric. A generalization of the

result to symmetric functions remains open. We also show that any threshold circuit of energy $e = 1$ needs at least 2^{n-1} threshold gates to compute PARITY of n variables.

References

- [1] D. Attwell and S. B. Laughlin. An energy budget for signaling in the gray matter of the brain. *Journal of Cerebral Blood Flow and Metabolism*, 21:1133–1145, 2001.
- [2] P. Földiák. Sparse coding in the primate cortex. *The Handbook of Brain Theory and Neural Networks*, 1:1064–1068, 2003.
- [3] S. B. Laughlin and T. J. Sejnowski. Communication in neuronal networks. *Science*, 301(5641):1870–1874, September 2003.
- [4] P. Lennie. The cost of cortical computation. *Current Biology*, 13:493–497, 2003.
- [5] M. Minsky and S. Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, Cambridge, MA, 1988.
- [6] B. A. Olshausen and D. J. Field. Sparse coding of sensory inputs. *Current Opinion in Neuro Biology*, 14:481–487, 2004.
- [7] I. Parberry. *Circuit Complexity and Neural Networks*. MIT Press, Cambridge, MA, 1994.
- [8] J. Sima and P. Orponen. General-purpose computation with neural networks: A survey of complexity theoretic results. *Neural Computation*, 15:2727–2778, 2003.
- [9] K. Y. Siu, V. Roychowdhury, and T. Kailath. *Discrete Neural Computation; A Theoretical Foundation*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1995.
- [10] K. Uchizawa, R. Douglas, and W. Maass. On the computational power of threshold circuits with sparse activity. *Neural Computation*, 18(12):2994–3008, 2006.
- [11] K. Uchizawa, T. Nishizeki, and E. Takimoto. Energy complexity and depth of threshold circuits. In *Proceedings of the 17th International Symposium on Fundamentals of Computation Theory*, Springer Lect. Notes in Computer Science 5699, pp. 335–345, 2009.
- [12] K. Uchizawa and E. Takimoto. Exponential lower bounds on the size of constant-depth threshold circuits with small energy complexity. *Theoretical Computer Science*, 407(1-3):474–487, 2008.
- [13] K. Uchizawa, E. Takimoto, and T. Nishizeki. Size and energy of threshold circuits computing mod functions. In *Proceedings of the 34th Int. Symp. on Mathematical Foundations of Computer Science*, Aug. 24–28, 2009, High Tatras, Slovakia, Springer Lect. Notes in Computer Science 5734, pp. 724–735, 2009.
- [14] W. E. Vinje and J. L. Gallant. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456):1273–1276, 2000.

Appendix: Proof of Theorem 3

Let C be a threshold circuit of size s and energy $e = 1$ that computes PARITY of n variables. We denote by g_1, g_2, \dots, g_s the threshold gates in C , and let g_s be the top gate of C . Let $X_0 = \{\mathbf{z} \in \{0, 1\}^n \mid |\mathbf{z}| \text{ is even}\}$, and n_0 be the cardinality of X_0 . Clearly, $n_0 = 2^{n-1}$. We prove that $s \geq n_0 = 2^{n-1}$ as follows.

For the sake of contradiction, assume that $s \leq n_0 - 1$. Since the top gate g_s outputs zero for any input $\mathbf{z} \in X_0$, we have, for each input $\mathbf{z} \in X_0$, either exactly one of g_1, g_2, \dots, g_{s-1} outputs one or none of the gates outputs one. Since $s \leq n_0 - 1$, the pigeon hole principle implies that there exists a pair of inputs $\mathbf{x} = (x_1, x_2, \dots, x_n), \mathbf{y} = (y_1, y_2, \dots, y_n) \in X_0$ that satisfies one of the following conditions:

- (i) there exists only an index k , $1 \leq k \leq s - 1$, such that the gate g_k outputs one for each of the inputs \mathbf{x} and \mathbf{y} ;
- (ii) none of the gates g_1, g_2, \dots, g_s outputs one for each of the inputs \mathbf{x} and \mathbf{y} .

For each of (i) and (ii), we derive a contradiction as follows.

We first consider (i). Let the gate g_k have weights w_1, w_2, \dots, w_n for the n input variables and a threshold t . Since only the gate g_k outputs one for each of \mathbf{x} and \mathbf{y} , we clearly have

$$\sum_{i=1}^n w_i x_i - t \geq 0 \text{ and } \sum_{i=1}^n w_i y_i - t \geq 0.$$

Thus,

$$\sum_{i=1}^n w_i x_i + \sum_{i=1}^n w_i y_i - 2t \geq 0. \quad (22)$$

Since $\mathbf{x} \neq \mathbf{y}$, there exists an index j such that $x_j \neq y_j$. Consider a pair of inputs \mathbf{x}' and \mathbf{y}' obtained from \mathbf{x} and \mathbf{y} by exchanging the j th components of \mathbf{x} for that of \mathbf{y} :

$$\mathbf{x}' = (x_1, x_2, \dots, x_{j-1}, y_j, x_{j+1}, \dots, x_n) \quad (23)$$

and

$$\mathbf{y}' = (y_1, y_2, \dots, y_{j-1}, x_j, y_{j+1}, \dots, y_n) \quad (24)$$

Both $|\mathbf{x}'|$ and $|\mathbf{y}'|$ are even and we have either $0 = x_j \neq y_j = 1$ or $1 = x_j \neq y_j = 0$, and hence $|\mathbf{x}'|$ and $|\mathbf{y}'|$ are odd. Thus, only the top gate g_s outputs one for each of \mathbf{x}' and \mathbf{y}' , and consequently g_k outputs zero for each of \mathbf{x}' and \mathbf{y}' , which implies that

$$\sum_{i=1}^n w_i x_i - w_j x_j + w_j y_j - t < 0 \text{ and } \sum_{i=1}^n w_i y_i - w_j y_j + w_j x_j - t < 0.$$

Thus,

$$\sum_{i=1}^n w_i x_i + \sum_{i=1}^n w_i y_i - 2t < 0. \quad (25)$$

We obtain a contradiction from Eqs. (22) and (25)

We next consider (ii), and derive a contradiction in a similar way to (i). Let the top gate g_s have weights w_1, w_2, \dots, w_n for the n input variables and a threshold t . Since none of the gates outputs one for \mathbf{x} and \mathbf{y} , we clearly have

$$\sum_{i=1}^n w_i x_i - t < 0 \text{ and } \sum_{i=1}^n w_i y_i - t < 0.$$

Thus,

$$\sum_{i=1}^n w_i x_i + \sum_{i=1}^n w_i y_i - 2t < 0. \quad (26)$$

Let j be an index such that $x_j \neq y_j$, then consider a pair of inputs \mathbf{x}' and \mathbf{y}' obtained from \mathbf{x} and \mathbf{y} by switching the j th components of the inputs as in Eqs (23) and (24). Clearly, $|\mathbf{x}'|$ and $|\mathbf{y}'|$ are both odd. Thus, the top gate g_s outputs one for each of \mathbf{x}' and \mathbf{y}' , which implies that

$$\sum_{i=1}^n w_i x_i - w_j x_j + w_j y_j - t \geq 0 \text{ and } \sum_{i=1}^n w_i y_i - w_j y_j + w_j x_j - t \geq 0.$$

Thus,

$$\sum_{i=1}^n w_i x_i + \sum_{i=1}^n w_i y_i - 2t \geq 0. \quad (27)$$

We obtain a contradiction from Eqs. (26) and (27) □