On Algebraic Constructions of Neural Networks with Small Weights

Kordag Mehmet Kilic, **Jin Sima** and **Jehoshua Bruck**

Electrical Engineering, California Institute of Technology, USA, {`kkilic,jsima,bruck`}`@caltech.edu`

Abstract—Neural gates compute functions based on weighted sums of the input variables. The expressive power of neural gates (number of distinct functions it can compute) depends on the weight sizes and, in general, large weights (exponential in the number of inputs) are required. Studying the trade-offs among the weight sizes, circuit size and depth is a well-studied topic both in circuit complexity theory and the practice of neural computation. We propose a new approach for studying these complexity trade-offs by considering a related algebraic framework. Specifically, given a single linear equation with arbitrary coefficients, we would like to express it using a system of linear equations with smaller (even constant) coefficients. The techniques we developed are based on Siegel's Lemma for the bounds, anti-concentration inequalities for the existential results and extensions of Sylvester-type Hadamard matrices for the constructions.

We explicitly construct a constant weight, optimal size matrix to compute the EQUALITY function (checking if two integers expressed in binary are equal). Computing EQUALITY with a single linear equation requires exponentially large weights. In addition, we prove the existence of the best-known weight size (linear) matrices to compute the COMPARISON function (comparing between two integers expressed in binary). In the context of the circuit complexity theory, our results improve the upper bounds on the weight sizes for the best-known circuit sizes for EQUALITY and COMPARISON.

I. INTRODUCTION

An n-input Boolean function is a mapping from $\{0,1\}^n$ to $\{0,1\}$. In other words, it is a partitioning of n-bit binary vectors into two sets with labels 0 and 1. In general, we can use systems of linear equations as descriptive models of these two sets of binary vectors. For example, the solution set of the equation $\sum_{i=1}^{n} x_i = k$ is the n-bit binary vectors $X = (x_1, \ldots, x_n)$ where each $x_i \in \{0,1\}$ and k is the number of 1s in the vectors. We can ask three important questions: How expressive can a single linear equation be? How many equations do we need to describe a Boolean function? Could we simulate a single equation by a system of equations with smaller integer weights?

Let us begin with an example: 3-input PARITY function where we label binary vectors with odd number of 1s as 1. We can write it in the following form:

$$\text{PARITY}(X) = \mathbb{1}\Big\{(2^2 x_3 + 2^1 x_2 + 2^0 x_1) \in \{1,2,4,7\}\Big\} \quad (1)$$

where $\mathbb{1}\{.\}$ is the indicator function with outputs 0 or 1. We express the binary vectors as integers by using binary expansions. Thus, it can be shown that if the weights are exponentially large in n, we can express all Boolean functions in this form.

Now, suppose that we are only allowed to use a single equality check in an indicator function. Considering the 3-input PARITY function, we can simply obtain

$$\begin{bmatrix} 2^0 & 2^1 & 2^2 \\ 2^0 & 2^1 & 2^2 \\ 2^0 & 2^1 & 2^2 \\ 2^0 & 2^1 & 2^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 7 \end{bmatrix} \quad (2)$$

None of the above equations can be satisfied if X is labeled as 0. Conversely, if X satisfies one of the above equations, we can label it as 1. For an arbitrary Boolean function of n inputs, if we list every integer associated with vectors labeled as 1, the number of rows may become exponentially large in n. Nevertheless, in this fashion, we can compute this function by the following system of equations using smaller weights.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad (3)$$

Not only there is a simplification on the number of equations, but also the weights are reduced to smaller sizes. This phenomenon motivates the following question with more emphasis: For which Boolean functions could we obtain such simplifications in the number of equations and weight sizes? For PARITY, this simplification is possible because it is a *symmetric* Boolean function, i.e., the output depends only on the number of 1s of the input X. We are particularly interested in such simplifications from large weights to small weights for a class of Boolean functions called *threshold functions*. Note that we use the word "large" to refer to exponentially large quantities in n and the word "small" to refer to polynomially large quantities (including $O(n^0) = O(1)$) in n.

A. Threshold Functions and Neural Networks

Threshold functions are commonly treated functions in Boolean analysis and machine learning as they form the basis of neural networks. Threshold functions compute a weighted summation of binary inputs and feed it to a *threshold* or *equality* check. If this sum is fed to the former, we call the functions *linear threshold functions* (see (4)). If it is fed to the latter, we call them *exact threshold functions* (see (5)). We

can write an n-input threshold function using the indicator function where w_is are integer weights and b is a *bias* term.

$$f_{\mathcal{LT}}(X) = \mathbb{1}\left\{\sum_{i=1}^{n} w_i x_i \geq b\right\} \qquad (4)$$

$$f_{\mathcal{E}}(X) = \mathbb{1}\left\{\sum_{i=1}^{n} w_i x_i = b\right\} \qquad (5)$$

A device computing the corresponding threshold function is called a *gate* of that type. To illustrate the concept, we define COMPARISON (denoted by COMP) function which computes whether an n-bit integer X is greater than another n-bit integer Y. The exact counterpart of it is defined as the EQUALITY (denoted by EQ) function which checks if two n-bit integers are equal (see Figure 1). For example, we can write the EQ function in the following form.

$$\text{EQ}(X, Y) = \mathbb{1}\left\{\sum_{i=1}^{n} 2^{i-1}(x_i - y_i) = 0\right\} \qquad (6)$$

Fig. 2. An example of a weight transformation for a single gate (in black) to construct constant weight circuits. Different gates are colored in red, blue, and violet and depending on the weight size, each gate is replicated a number of times. In this example, each weight in this construction is at most 3.
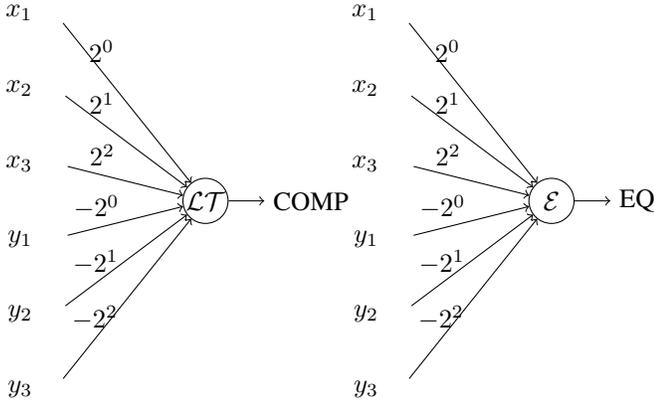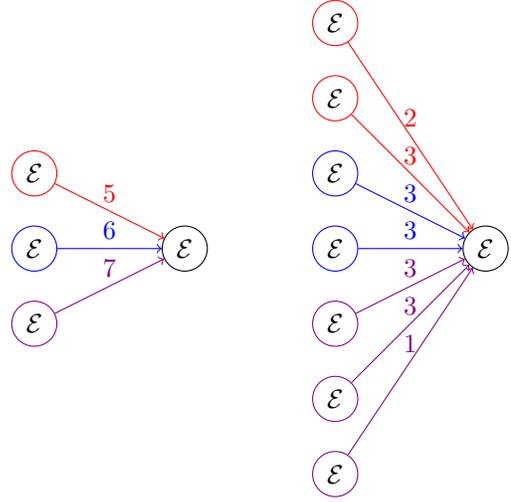

Fig. 1. The 3-input COMP and EQ functions for integers X and Y computed by linear threshold and exact threshold gates. A gate with an $\mathcal{LT}$(or $\mathcal{E}$) inside is a linear (or exact) threshold gate. More explicitly, we can write $\text{COMP}(X, Y) = \mathbb{1}\{4x_3 + 2x_2 + x_1 \geq 4y_3 + 2y_2 + y_1\}$ and $\text{EQ}(X, Y) = \mathbb{1}\{4x_3 + 2x_2 + x_1 = 4y_3 + 2y_2 + y_1\}$

In general, it is proven that the weights of a threshold function can be represented by $O(n \log n)$-bits and this is tight [1, 3, 11, 19]. However, it is possible to construct "small" weight *threshold circuits* to compute any threshold function [2, 8, 12, 23]. This transformation from a circuit of depth d with exponentially large weights in n to another circuit with polynomially large weights in n is typically within a constant factor of depth (e.g. $d + 1$ or $3d + 3$ depending on the context) [8, 25]. For instance, such a transformation would simply follow if we can replace any "large" weight threshold function with "small" weight depth-2 circuits so that the new depth becomes $2d$.

It is possible to reduce polynomial size weights into constant weights by replicating the gates that is fed to the top gate recursively (see Figure 2). Nevertheless, this would inevitably introduce a polynomial size blow-up in the circuit size. We

emphasize that our focus is to achieve this weight size reduction from polynomial weights to constant weights with at most a constant size blow-up in the circuit size.

For neural networks and learning tasks, the weights are typically finite precision real numbers. To make the computations faster and more power efficient, the weights can be quantized to small integers with a loss in the accuracy of the network output [13, 14]. In practice, given that the neural circuit is large in size and deep in depth, this loss in the accuracy might be tolerated. We are interested in the amount of trade-off in the increase of size and depth while using as small weights as possible. More specifically, our goal is to provide insight to the computation of single threshold functions with large weights using threshold circuits with small weights by relating theoretical upper and lower bounds to the best known constructions.

In this manner, for the ternary weight case ($\{-1, 0, 1\}$ as in [17]), we give an explicit and optimal size circuit construction for the EQ function using depth-2 circuits. This optimality is guaranteed by achieving the theoretical lower bounds asymptotically up to vanishing terms [16, 20]. We also prove an existential result on the COMP constructions to reduce the weight size on the best known constructions. It is not known if constant weight constructions exist without a polynomial blow-up in the circuit size and an increase in the depth for arbitrary threshold functions.

B. Bijective Mappings from Finite Fields to Integers

It seems that choosing powers of two as the weight set is important. In fact, we can expand any arbitrary weight in binary and use powers of two as a fixed weight set for any threshold function [16]. This weight set is a choice of convenience and a simple solution for the following question: How small could the elements of $\mathcal{W} = \{w_1, w_2, \cdots, w_n\}$

be if $\mathcal{W}$ has all distinct subset sums (DSS)? If the weights satisfy this property, called the **DSS property**, they can define a bijection between $\{0,1\}^n$ and integers. Erdős conjectured that the largest weight $w \in \mathcal{W}$ is upper bounded by $c_0 2^n$ for some $c_0 > 0$ and therefore, choosing powers of two as the weight set is asymptotically optimal. The best known result for such weight sets yields $0.22002 \cdot 2^n$ and currently, the best lower bound is $\Omega(2^n/\sqrt{n})$ [5, 7, 9].

Now, let us consider the following linear equation where the weights are fixed to the ascending powers of two but x_is are not necessarily binary. We denote the powers of two by the vector w_b.

$$w_b^T x = \sum_{i=1}^n 2^{i-1} x_i = 0 \tag{7}$$

As the weights of w_b define a bijection between n-bit binary vectors and integers, $w_b^T x = 0$ does not admit a non-trivial solution for the alphabet $\{-1, 0, 1\}^n$. This is a necessary and sufficient condition to compute the EQ function given in (6). We extend this property to m many rows to define EQ *matrices* which give a bijection between $\{0,1\}^n$ and $\mathbb{Z}^m$. Thus, an EQ matrix can be used to compute the EQ function in (6) and our goal is to use smaller weight sizes in the matrix.

Definition 1. *A matrix $A \in \mathbb{Z}^{m \times n}$ is an* EQ *matrix if the homogeneous system $Ax = 0$ has no non-trivial solutions in $\{-1, 0, 1\}^n$.*

Let $A \in \mathbb{Z}^{m \times n}$ be an EQ matrix with the weight constraint $W \in \mathbb{Z}$ such that $|a_{ij}| \leq W$ for all i, j and let R denote the *rate* of the matrix A, which is n/m. It is clear that any full-rank square matrix is an EQ matrix with $R = 1$. Given any W, how large can this R be? For the maximal rate, a necessary condition can be proven by Siegel's Lemma [22].

Lemma 1 (Siegel's Lemma (modified)). *Consider any integer matrix $A \in \mathbb{Z}^{m \times n}$ with $m < n$ and $|a_{ij}| \leq W$ for all i, j and some integer W. Then, $Ax = 0$ has a non-trivial solution for an integer vector $x \in \mathbb{Z}^n$ such that $||x||_\infty \leq (\sqrt{n}W)^{\frac{m}{n-m}}$.*

It is shown that if $m = o(n/\log nW)$, then any $A \in \mathbb{Z}^{m \times n}$ with weight constraint W admits a non-trivial solution in $\{-1, 0, 1\}^n$ and cannot be an EQ matrix, i.e., $R = O(\log nW)$ is tight [16]. A similar result can be obtained by the matrix generalizations of Erdős' Distinct Subset Sum problem [6].

When $m = O(n/\log nW)$, the story is different. If $m = O(n/\log n)$, there exists a matrix $A \in \{-1, 1\}^{m \times n}$ such that every non-trivial solution of $Ax = 0$ satisfies $\max_j |x_j| \geq c_0\sqrt{n}^{\frac{m}{n-m}}$ for a positive constant c_0. This is given by Beck's Converse Theorem on Siegel's Lemma [4].

For an explicit construction, it is possible to achieve the optimal rate $R = O(\log nW)$ if we allow $W = poly(n)$. This can be done by the Chinese Remainder Theorem (CRT) and the Prime Number Theorem (PNT) [2, 12, 16]. It is known that CRT can be used to define *residue codes* [18]. For an integer x and modulo base p, we denote the modulo operation by $[x]_p$, which maps the integer to a value in $\{0, ..., p-1\}$. Suppose $0 \leq Z < 2^n$ for an integer Z. One can encode this

integer by the m-tuple $(d_1, d_2, \cdots, d_m)$ where $[Z]_{p_i} = d_i$ for a prime number p_i. Since we can also encode Z by its binary expansion, the CRT gives a bijection between $\mathbb{Z}^m$ and $\{0,1\}^n$ as long as $p_1 \cdots p_m > 2^n$. By taking modulo p_i of Equation (7), we can obtain the following matrix, defined as a *CRT matrix*:

$$\begin{bmatrix} [2^0]_3 & [2^1]_3 & [2^2]_3 & [2^4]_3 & [2^5]_3 & [2^6]_3 & [2^7]_3 \\ [2^0]_5 & [2^1]_5 & [2^2]_5 & [2^4]_5 & [2^5]_5 & [2^6]_5 & [2^7]_5 \\ [2^0]_7 & [2^1]_7 & [2^2]_7 & [2^4]_7 & [2^5]_7 & [2^6]_7 & [2^7]_7 \\ [2^0]_{11} & [2^1]_{11} & [2^2]_{11} & [2^4]_{11} & [2^5]_{11} & [2^6]_{11} & [2^7]_{11} \end{bmatrix} \tag{8}$$

$$= \begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 \\ 1 & 2 & 4 & 3 & 1 & 2 & 4 & 3 \\ 1 & 2 & 4 & 1 & 2 & 4 & 1 & 2 \\ 1 & 2 & 4 & 8 & 5 & 10 & 9 & 7 \end{bmatrix}_{4 \times 8} \tag{9}$$

We have $Z < 256$ since $n = 8$ and $3 \cdot 5 \cdot 7 \cdot 11 = 1155$. Therefore, this CRT matrix is an EQ matrix. In general, by the PNT, one needs $O(n/\log n)$ many rows to ensure that $p_1 \cdots p_m > 2^n$. Moreover, W is bounded by the maximum prime size p_m, which is $O(n)$ again by the PNT.

However, it is known that constant weight EQ matrices with asymptotically optimal rate exist by Beck's Converse Theorem on Siegel's Lemma [4, 16]. In this paper, we give an explicit construction where $W = 1$ and asymptotic efficiency in rate is achieved up to vanishing terms. It is in fact an extension of Sylvester-type Hadamard matrices.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 & -1 & 0 & 0 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 \\ 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}_{4 \times 8} \tag{10}$$

One can verify that the matrix in (10) is an EQ matrix and its rate is twice the trivial rate being the same in (8). Therefore, due to the optimality in the weight size, this construction can replace the CRT matrices in the constructions of EQ matrices.

We can also focus on q-ary representation of integers in a similar fashion by extending the definition of EQ matrices to this setting. In our analysis, we always treat q as a constant value.

Definition 2. *A matrix $A \in \mathbb{Z}^{m \times n}$ is an* EQ$_q$ *matrix if the homogeneous system $Ax = 0$ has no non-trivial solutions in $\{-q+1, \ldots, q-1\}^n$.*

If $q = 2$, then we drop q from the notation and say that the matrix is an EQ matrix. For the EQ$_q$ matrices, the optimal rate given by Siegel's Lemma is still $R = O(\log nW)$ and constant weight constructions exist. We give an extension of our construction to EQ$_q$ matrices where $W = 1$ and asymptotic efficiency in rate is achieved up to constant terms.

C. Maximum Distance Separable Extensions of EQ Matrices

Residue codes are treated as Maximum Distance Separable (MDS) codes because one can extend the CRT matrix by adding more prime numbers to the matrix (without increasing

n) so that the resulting integer code achieves the Singleton bound in Hamming distance [24]. However, we do not say a CRT matrix is an *MDS matrix* as this refers to another concept.

Definition 3. *An integer matrix $A \in \mathbb{Z}^{m \times n}$ ($m \leq n$) is* MDS *if and only if no $m \times m$ submatrix B is singular.*

Definition 4. *An integer matrix $A \in \mathbb{Z}^{rm \times n}$ is* MDS *for q-ary bijections with MDS rate r and EQ rate $R = n/m$ if and only if every $m \times n$ submatrix B is an* EQ_q *matrix.*

Because Definition 4 considers solutions over a **restricted** alphabet, we denote such matrices as $RMDS_q$. Remarkably, as $q \to \infty$, both MDS definitions become the same. Similar to the EQ_q definition, we drop the q from the notation when $q = 2$. A CRT matrix is not MDS, however, it can be $RMDS_q$.

We can demonstrate the difference of both MDS definitions by the following matrix. This matrix is an RMDS matrix with EQ rate 2 and MDS rate $5/4$ because any 4×8 submatrix is an EQ matrix. This is in fact the same matrix in (8) with an additional row with entries $[2^{i-1}]_{13}$ for $i \in \{1, ..., 8\}$.

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 \\ 1 & 2 & 4 & 3 & 1 & 2 & 4 & 3 \\ 1 & 2 & 4 & 1 & 2 & 4 & 1 & 2 \\ 1 & 2 & 4 & 8 & 5 & 10 & 9 & 7 \\ 1 & 2 & 4 & 8 & 3 & 6 & 12 & 11 \end{bmatrix}_{5 \times 8} \quad (11)$$

Here, the determinant of the 5×5 submatrix given by the first five columns is 0. Thus, this matrix is not an MDS matrix.

In this work, we are interested in $RMDS_q$ matrices. We show that for a constant weight $RMDS_q$ matrix, the MDS rate bound $r = O(1)$ is tight. We also provide an existence result for such $RMDS_q$ matrices where the weight size is bounded by $O(r)$ given that the EQ rate is $O(\log n)$.

The following is a summary of our contributions in this paper.

- In Section II, we explicitly give a rate-efficient EQ matrix construction with constant entries $\{-1, 0, 1\}$ where the optimality is guaranteed up to vanishing terms. This solves an open problem in [16] and [20].
- In Section III, we prove that the MDS rate r of an $RMDS_q$ matrix with entries from an alphabet $\mathcal{Q}$ with cardinality k should satisfy $r \leq k^{k+1}$. Therefore, constant weight $RMDS_q$ matrices can at most achieve $r = O(1)$.
- In Section III, provide an existence result for $RMDS_q$ matrices given that $W = O(r)$ and the optimal EQ rate $O(\log n)$. In contrast, the best known results give $W = O(rn)$ with the optimal EQ rate $O(\log n)$.
- In Section IV, we apply our results to Circuit Complexity Theory to obtain better weight sizes with asymptotically **no trade-off** in the circuit size for the depth-2 EQ and COMP constructions, as shown in the Table I.

II. RATE-EFFICIENT CONSTRUCTIONS WITH CONSTANT ALPHABET SIZE

The $m \times n$ EQ matrix construction we give here is based on Sylvester's construction of Hadamard matrices. It is an

TABLE I
RESULTS FOR THE DEPTH-2 CIRCUIT CONSTRUCTIONS

Function	This Work		Previous Works	
	Weight Size	Constructive	Weight Size	Constructive
EQ	$O(1)$	Yes	$O(1)$[16]	No
			$O(n)$[20]	Yes
COMP	$O(n)$	No	$O(n^2)$[2]	Yes

extension of it to achieve higher rates with the trade-off that the matrix is no longer full-rank.

Theorem 1. *Suppose we are given an EQ matrix $A_0 \in \{-1, 0, 1\}^{m_0 \times n_0}$. At iteration k, we construct the following matrix A_k:*

$$A_k = \begin{bmatrix} A_{k-1} & A_{k-1} & I_{m_{k-1}} \\ A_{k-1} & -A_{k-1} & 0 \end{bmatrix} \quad (12)$$

A_k is an EQ matrix with $m_k = 2^k m_0$, $n_k = 2^k n_0(\frac{k}{2}\frac{m_0}{n_0} + 1)$ for any integer $k \geq 0$.

Proof. We will apply induction. The case $k = 0$ is trivial by assumption. For the system $A_k x = z$, let us partition the vector x and z in the following way:

$$\begin{bmatrix} A_{k-1} & A_{k-1} & I_{m_{k-1}} \\ A_{k-1} & -A_{k-1} & 0 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ x^{(3)} \end{bmatrix} = \begin{bmatrix} z^{(1)} \\ z^{(2)} \end{bmatrix} \quad (13)$$

Then, setting $z = 0$, we have $A_{k-1}x^{(1)} = A_{k-1}x^{(2)}$ by the second row block. Hence, the first row block of the construction implies $2A_{k-1}x^{(1)} + x^{(3)} = 0$. Each entry of $x^{(3)}$ is either 0 or a multiple of 2. Since $x_i \in \{-1, 0, 1\}$, we see that $x^{(3)} = 0$. Then, we obtain $A_{k-1}x^{(1)} = A_{k-1}x^{(2)} = 0$. Applying the induction hypothesis on $A_{k-1}x^{(1)}$ and $A_{k-1}x^{(2)}$, we see that $A_k x = 0$ admits a unique trivial solution in $\{-1, 0, 1\}^{n_k}$.

To see that $m_k = 2^k m_0$ is not difficult. For n_k, we can apply induction. $\square$

To construct the matrix given in (10), we can use $A_0 = I_1 = \begin{bmatrix} 1 \end{bmatrix}$, which is a trivial rate EQ matrix, and take $k = 2$. By rearrangement of the columns, we also see that there is a 4×4 Sylvester-type Hadamard matrix in this matrix.

For the sake of completeness, we will revisit Siegel's Lemma to prove the best possible rate one can obtain for an EQ matrix with weights $\{-1, 0, 1\}$ as similarly done in [16]. We note that the following lemma gives the best possible rate upper bound to the best of our knowledge and our construction asymptotically shows the sharpness of Siegel's Lemma similar to Beck's result.

Lemma 2. *For any $\{-1, 0, 1\}^{m \times n}$ EQ matrix, $R = \frac{n}{m} \leq \frac{1}{2}\log n + 1$.*

Proof. By Siegel's Lemma, we know that for a matrix $A \in \mathbb{Z}^{m \times n}$, the homogeneous system $Ax = 0$ attains a non-trivial solution in $\{-1, 0, 1\}^n$ when

$$\|x\|_\infty \leq (\sqrt{n}W)^{\frac{m}{n-m}} \leq 2^{1-\epsilon} \quad (14)$$

for some $\epsilon > 0$. Then, since $W = 1$, we deduce that $R = \frac{n}{m} \geq \frac{1}{2(1-\epsilon)} \log n + 1$. Obviously, an EQ matrix cannot obtain such a rate. Taking $\epsilon \to 0$, we conclude the best upper bound is $R \leq \frac{1}{2} \log n + 1$. $\qquad\square$

Using Lemma 2 for our construction with $A_0 = \begin{bmatrix} 1 \end{bmatrix}$, we compute the upper bound on the rate and the real rate

$$R_{upper} = \frac{k + 1 + \log (k + 2)}{2}, \quad R_{constr} = \frac{k}{2} + 1 \quad (15)$$

$$\frac{R_{upper}}{R_{constr}} = 1 + \frac{\log (k + 2) - 1}{k + 2} \implies \frac{R_{upper}}{R_{constr}} \sim 1 \quad (16)$$

which concludes that the construction is **optimal** in rate up to vanishing terms. By deleting columns, we can generalize the result to any $n \in \mathbb{Z}$ to achieve optimality up to a constant factor of 2.

We conjecture that one can extend the columns of any Hadamard matrix to obtain an EQ matrix with entries $\{-1, 0, 1\}$ achieving rate optimality up to vanishing terms. In this case, the Hadamard conjecture implies a rich set of EQ matrix constructions.

Given $Ax = z \in \mathbb{Z}^m$, we note that there is a linear time decoding algorithm to find $x \in \{0, 1\}^n$ uniquely, given in the appendix. This construction can also be generalized to EQ_q matrices with the same proof idea. In this case, the rate is optimal up to a factor of q.

Theorem 2. *Suppose we are given an EQ_q matrix $A_0 \in \{-1, 0, 1\}^{m_0 \times n_0}$. At iteration k, we construct the following matrix A_k:*

$$\begin{bmatrix} A_{k-1} & A_{k-1} & A_{k-1} & \cdots & A_{k-1} & I_{m_{k-1}} \\ A_{k-1} & -A_{k-1} & 0 & \cdots & 0 & 0 \\ 0 & A_{k-1} & -A_{k-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -A_{k-1} & 0 \end{bmatrix} \quad (17)$$

A_k is an EQ_q matrix with $m_k = q^k m_0$, $n_k = q^k n_0 (\frac{k}{q} \frac{m_0}{n_0} + 1)$ for any integer $k \geq 0$.

III. BOUNDS ON THE RATE AND THE WEIGHT SIZE OF RMDS_q MATRICES

Similar to a CRT-based RMDS matrix, is there a way to extend the EQ matrix given in Section II without a large trade-off in the weight size? We give an upper bound on the MDS rate r based on the alphabet size of an RMDS_q matrix.

Theorem 3. *An RMDS_q matrix $A \in \mathbb{Z}^{rm \times n}$ with entries from $\mathcal{Q} = \{q_1, ..., q_k\}$ satisfies $r \leq k^{k+1}$ given that $n > k$.*

Proof. The proof is based on a simple counting argument. We rearrange the rows of a matrix in blocks B_i for $i \in \{1, \cdots, k^{k+1}\}$ assuming $n > k$. Here, each block contains rows starting with a prefix of length $k+1$ in the lexicographical order of the indices, i.e.,

$$A = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_{k^{k+1}} \end{bmatrix} = \begin{bmatrix} q_1 & q_1 & \cdots & q_1 & \cdots \\ q_1 & q_1 & \cdots & q_2 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ q_k & q_k & \cdots & q_k & \cdots \end{bmatrix} \quad (18)$$

For instance, the block B_1 contains the rows starting with the prefix $\begin{bmatrix} q_1 & q_1 & \cdots & q_1 \end{bmatrix}_{1 \times k+1}$. It is easy to see that there is a vector $x \in \{-1, 0, 1\}^n$ that will make at least one of the $B_i x = 0$ because it is guaranteed that the $(k + 1)$-th column will be equal to the one of the preceding elements by Pigeonhole Principle.

Since the matrix is MDS, any m row selections should be an EQ matrix. Therefore, any B_i should not contain more than m rows. Again, by Pigeonhole Principle, $rm \leq k^{k+1}m$ and consequently, $r \leq k^{k+1}$. $\qquad\square$

The condition that $n > k$ is trivial in the sense that if the weights are constant, then k is a constant. Therefore, the MDS rate can only be a constant due to Theorem 3.

Corollary 3.1. *An RMDS_q matrix $A \in \mathbb{Z}^{rm \times n}$ weight size $W = O(1)$ can at most achieve the MDS rate $r = O(1)$.*

In Section II, we saw that CRT-based EQ matrix constructions are not optimal in terms of the weight size. For RMDS_q matrices, we now show that the weight size can be reduced by a factor of n. The idea is to use the probabilistic method on the existence of RMDS_q matrices.

Lemma 3. *Suppose that $a \in \{-W, \ldots, W\}^n$ is uniformly distributed and $x_i \in \{-q+1, \ldots, q-1\} \setminus 0$ for $i \in \{1, \ldots, n\}$ are fixed where q is a constant. Then, for some constant C,*

$$Pr(a^T x = 0) \leq \frac{C}{\sqrt{nW}} \quad (19)$$

Proof. We start with a statement of the Berry-Esseen Theorem.

Lemma 4 (Berry-Esseen Theorem). *Let $X_1, \ldots, X_n$ be independent centered random variables with finite third moments $\mathbb{E}[|X_i^3|] = \rho_i$ and let $\sigma^2 = \sum_{i=1}^{n} \mathbb{E}[X_i^2]$. Then, for any $t > 0$,*

$$\left| Pr\left(\sum_{i=1}^{n} X_i \leq t \right) - \Phi(t) \right| \leq C\sigma^{-3} \sum_{i=1}^{n} \rho_i \quad (20)$$

where C is an absolute constant and $\Phi(t)$ is the cumulative distribution function of $\mathcal{N}(0, \sigma^2)$.

Since the density of a normal random variable is uniformly bounded by $1/\sqrt{2\pi\sigma^2}$, we obtain the following.

$$Pr\left(\left| \sum_{i=1}^{n} X_i \right| \leq t \right) \leq \frac{2t}{\sqrt{2\pi\sigma^2}} + 2C\sigma^{-3} \sum_{i=1}^{n} \rho_i \quad (21)$$

Let $a^{(n-1)}$ denote the $(a_1, \ldots, a_{n-1})$ and similarly, let $x^{(n-1)}$ denote $(x_1, \ldots, x_{n-1})$. By the Total Probability The-

orem, we have

$$Pr(a^T x = 0) = Pr\left(\frac{a^{(n-1)T} x^{(n-1)}}{|x_n|} \in \{-W, \ldots, W\}\right)$$

$$Pr\left(a^T x = 0 \middle| \frac{a^{(n-1)T} x^{(n-1)}}{|x_n|} \in \{-W, \ldots, W\}\right)$$

$$= Pr\left(\frac{a^{(n-1)T} x^{(n-1)}}{|x_n|} \in \{-W, \ldots, W\}\right) \frac{1}{2W+1} \tag{22}$$

where the last line follows from the fact that $Pr(a_n = k)$ for some $k \in \{-W, \ldots, W\}$ is $\frac{1}{2W+1}$. The other conditional probability term is 0 because $Pr(a_n = k)$ for any $k \notin \{-W, \ldots, W\}$ is 0. We will apply Berry-Esseen Theorem to find an upper bound on the first term.

We note that $\mathbb{E}[a_i^2 x_i^2] = x_i^2 \frac{(2W+1)^2-1}{12}$ and $\mathbb{E}[|a_i^3 x_i^3|] = |x_i|^3 \frac{2}{2W+1}\left(\frac{W(W+1)}{2}\right)^2$. Then,

$$Pr\left(\left|\frac{a^{(n-1)T} x^{(n-1)}}{|x_n|}\right| \leq W\right) \leq \frac{2W|x_n|}{\sqrt{2\pi}\sqrt{\frac{(2W+1)^2-1}{12} \sum_{i=1}^{n-1} x_i^2}}$$

$$+ C \frac{\frac{2}{2W+1}\left(\frac{W(W+1)}{2}\right)^2 \sum_{i=1}^{n-1} |x_i|^3}{\left(\frac{(2W+1)^2-1}{12}\right)^{\frac{3}{2}} \left(\sum_{i=1}^{n-1} x_i^2\right)^{\frac{3}{2}}} \tag{23}$$

One can check that the whole RHS is in the form of $\frac{C'}{\sqrt{n-1}}$ for a constant C' as we assume that q is a constant. We can bound $|x_i|$s in the numerator by $q-1$ and $|x_i|$s in the denominator by 1. The order of W terms are the same and we can use a limit argument to bound them as well. Therefore, for some $C'' > 0$,

$$Pr(a^T x = 0) \leq \frac{C'}{\sqrt{n-1}} \frac{1}{2W+1} \leq \frac{C''}{\sqrt{n}W} \tag{24}$$

$\square$

Lemma 3 is related to the Littlewood-Offord problem and anti-concentration inequalities are typically used in this framework [10, 21]. We specifically use Berry-Esseen Theorem. Building on this lemma and the union bound, we have the following theorem.

Theorem 4. *An* $RMDS_q$ *matrix* $M \in \mathbb{Z}^{rm \times n}$ *with entries in* $\{-W, \ldots, W\}$ *exists if* $m = \Omega(n/\log n)$ *and* $W = O(r)$.

Proof. Let A be any $m \times n$ submatrix of M. Then, by the union bound (as done in [15]), we have

$$Pr(\text{M is not RMDS}_q) = P \leq \binom{rm}{m} Pr(\text{A is not EQ}_q)$$

$$\leq r^m e^m Pr(\text{A is not EQ}_q) \tag{25}$$

We again use the union bound to sum over all $x \in \{-q+1, \ldots, q-1\}^n \setminus 0$ the event that $Ax = 0$ by counting non-zero entries in x by k. Also, notice the independence of the events

that $a_i^T x = 0$ for $i \in \{1, \ldots, m\}$. Let $x^{(k)}$ denote an arbitrary $x \in \{-q+1, \ldots, q-1\}^n \setminus 0$ with k non-zero entries. Then,

$$Pr(\text{A is not EQ}_q) \leq \sum_{x \in \{-q+1, \ldots, q-1\}^n \setminus 0} \prod_{i=1}^{m} Pr(a_i^T x = 0) \tag{26}$$

$$\leq \sum_{k=1}^{n} \binom{n}{k} (2(q-1))^k Pr(a^T x^{(k)} = 0)^m \tag{27}$$

We can use Lemma 3 to bound the $Pr(a^T x^{(k)} = 0)$ term. Therefore,

$$P \leq r^m e^m \sum_{k=1}^{n} \binom{n}{k} (2(q-1))^k \left(\frac{C}{\sqrt{k}W}\right)^m \tag{28}$$

$$= \sum_{k=1}^{T-1} \binom{n}{k} (2(q-1))^k \left(\frac{rC_1}{\sqrt{k}W}\right)^m$$

$$+ \sum_{k=T}^{n} \binom{n}{k} (2(q-1))^k \left(\frac{rC_1}{\sqrt{k}W}\right)^m \tag{29}$$

for some $C_1, T \in \mathbb{Z}$. We bound the first summation by using $\binom{n}{k}(2(q-1))^k \leq (n(q-1))^k$ and choosing $k = 1$ for the probability term. This gives a geometric sum from $k = 1$ to $k = T - 1$.

$$\sum_{k=1}^{T-1} (n(q-1))^k \left(\frac{rC_1}{W}\right)^m = \left(\frac{rC_1}{W}\right)^m \left(\frac{(n(q-1))^T - 1}{n(q-1) - 1}\right) \tag{30}$$

For the second term, we take the highest value term $\frac{rC_1}{\sqrt{T}W}$ and $\sum_{k=T}^{n} \binom{n}{k} 2(q-1)^k \leq \sum_{k=0}^{n} \binom{n}{k} 2(q-1)^k = (2(q-1)+1)^n$. Hence,

$$P \leq \left(\frac{rC_1}{W}\right)^m C_2(n(q-1))^T + (2(q-1)+1)^n \left(\frac{rC_1}{\sqrt{T}W}\right)^m \tag{31}$$

$$\leq 2^{C_3 T \log n(q-1) - m \log W/rC_1}$$

$$+ 2^{n \log(2(q-1)+1) - m \log(\sqrt{T}W/rC_1)} \tag{32}$$

We take $T = O(n^c)$ for some $0 < c < 1$. It is easy to see that if $m = \Omega(n/\log n)$ and $W = O(r)$, both terms vanish as n goes to the infinity.

$\square$

When $r = 1$, we prove an existential result on EQ matrices already known in [16]. We remark that the proof technique for Theorem 4 is not powerful enough to obtain non-trivial bounds on W when $m = 1$ to attack Erdős' conjecture on the DSS weight sets.

The CRT gives an explicit way to construct $RMDS_q \in \mathbb{Z}^{rm \times n}$ matrices. We obtain $p_{rm} = O(rm \log rm) = O(rn)$ given that $r = O(n^c)$ for some $c > 0$ and $m = O(n/\log n)$

by the PNT. Therefore, we have a factor of $O(n)$ weight size reduction in Theorem 4. However, modular arithmetical properties of the CRT do not reflect to $RMDS_q$ matrices. Therefore, an $RMDS_q$ matrix cannot replace a CRT matrix in general (see Appendix).

IV. APPLICATIONS OF THE ALGEBRAIC RESULTS TO NEURAL CIRCUITS

In this section, we will give EQ and COMP threshold circuit constructions. We note that in our analysis, the size of the bias term is ignored.

One can construct a depth-2 exact threshold circuit with small weights to compute the EQ function [16]. For the first layer, we select the weights for each exact threshold gate as the rows of the EQ matrix. Then, we connect the outputs of the first layer to the top gate, which just computes the m-input AND function (i.e. $\mathbb{1}\{z_1 + ... + z_m = m\}$ for $z_i \in \{0,1\}$). In Figure 3, we give an example of EQ constructions.
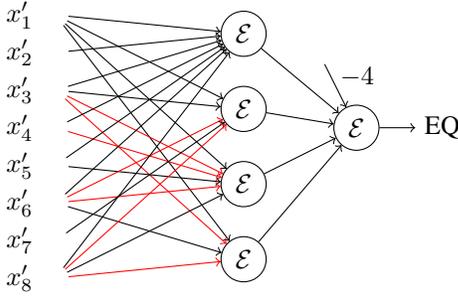

Fig. 3. An example of $EQ(X, Y)$ constructions with 8 $x'_i = x_i - y_i$ inputs (or 16 if x_is and y_is are counted separately) and 5 exact threshold gates (including the top gate). The black(or red) edges correspond to the edges with weight 1 (or -1).

We roughly follow the previous works for the construction of the COMP [2]. First, let us define $F^{(l)}(X, Y) = \sum_{i=l+1}^{n} 2^{i-l-1}(x_i - y_i)$ so that $\mathbb{1}\{F^{(l)}(X, Y) \geq 0\}$ is a $(n-l)$-bit COMP function for $l < n$ and $F^{(0)}(X, Y) = F(X, Y)$. We say that $X \geq Y$ when $F(X, Y) \geq 0$ and vice versa. We also denote by $X^{(l)}$ the most significant $(n-l)$-tuple of an n-tuple vector X. We have the following observation.

Lemma 5. *Let* $F^{(l)}(X, Y) = \sum_{i=l+1}^{n} 2^{i-l-1}(x_i - y_i)$ *and* $F^{(0)}(X, Y) = F(X, Y)$ *for* $X, Y \in \{0,1\}^n$. *Then,*

$$F(X, Y) > 0 \Leftrightarrow \exists l : F^{(l)}(X, Y) = 1 \quad (33)$$

$$F(X, Y) < 0 \Leftrightarrow \exists l : F^{(l)}(X, Y) = -1 \quad (34)$$

Proof. It is easy to see that if $X - Y = (0, 0, \ldots, 0, 1, \times, \ldots, \times)$ where the vector has a number of leading 0s and $\times$s denote any of $\{-1, 0, 1\}$, we see that $F(X, Y) > 0$ (this is called the *domination property*). Similarly, for $F(X, Y) < 0$, the vector $X - Y$ should have the form $(0, 0, \ldots, 0, -1, \times, \ldots, \times)$ and $F(X, Y) = 0$ if and only if $X - Y = (0, \ldots, 0)$. The converse holds similarly. $\square$

By searching for the $(n - l)$-tuple vectors $(X - Y)^{(l)} = (0, ..., 0, -1)$ for all $l \in \{0, ..., n - 1\}$, we can compute
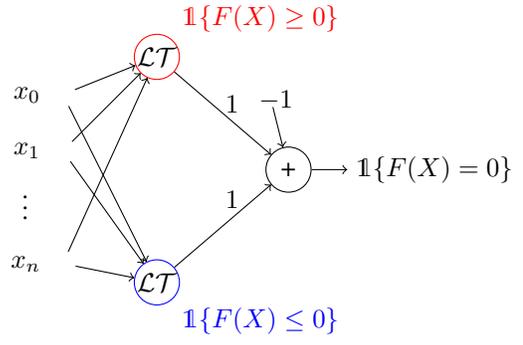

Fig. 4. A construction of an arbitrary exact threshold function ($\mathbb{1}\{F(X) = 0\}$) using two linear threshold functions ($\mathbb{1}\{F(X) \geq 0\}$ and $\mathbb{1}\{F(X) \leq 0\}$) and a summation node. This summation node can be removed if its output is connected to another gate due to linearity.

the COMP. We claim that if we have an $RMDS_q$ matrix $A \in \mathbb{Z}^{rm \times n}$, we can detect such vectors by solving $A^{(l)}(X - Y)^{(l)} = -a_{n-l}$ where $A^{(l)}$ is a truncated version of A with the first $n - l$ columns and a_{n-l} is the $(n - l)$-th column. Specifically, we obtain the following:

$$X < Y \Rightarrow \sum_{l=0}^{n-1} \mathbb{1}\{A^{(l)}(X - Y)^{(l)} = -a_{n-l}\} \geq rm \quad (35)$$

$$X \geq Y \Rightarrow \sum_{l=0}^{n-1} \mathbb{1}\{A^{(l)}(X - Y)^{(l)} = -a_{n-l}\} < n(m - 1) \quad (36)$$

Here, the indicator function works row-wise, i.e., we have the output vector $z \in \{0,1\}^{rmn}$ such that $z_{rml+i} = \mathbb{1}\{(A^{(l)}(X - Y)^{(l)})_i = -(a_{n-l})_i\}$ for $i \in \{1, \ldots, rm\}$ and $l \in \{0, \ldots, n-1\}$. We use an $RMDS_3$ matrix in the construction to map $\{-1, 0, 1\}^{n-l}$ vectors bijectively to integer vectors with large Hamming distance. Note that each exact threshold function can be replaced by two linear threshold functions and a summation layer (i.e. $\mathbb{1}\{F(X) = 0\} = \mathbb{1}\{F(X) \geq 0\} + \mathbb{1}\{-F(X) \geq 0\} - 1$) which can be absorbed to the top gate (see Figure 4) . This increases the circuit size only by a factor of two.

If $F(X, Y) < 0$, z_{rml+i} should be 1 for all $i \in \{1, \ldots, rm\}$ for some l by Lemma 5. For $F(X, Y) \geq 0$ and any l, the maximum number of 1s that can appear in z_{rml+i} is upper bounded by $m - 1$ because A is $RMDS_3$. Therefore, the maximum number of 1s that can appear in z is upper bounded by $n(m - 1)$. A sketch of the construction is given in Figure 5.

In order to make both cases separable, we choose $r = n$. At the second layer, the top gate is a MAJORITY gate ($\mathbb{1}\{\sum_{i=1}^{rmn} z_i < n(m - 1)\} = \mathbb{1}\{\sum_{i=1}^{rmn} -z_i \geq -n(m - 1) + 1\}$). By Theorem 4, there exists an $RMDS_3$ matrix with $m = O(n/\log n)$ and $W = O(n)$. Thus, there are $rmn + 1 = O(n^3/\log n)$ many gates in the circuit, which is the best known result. The same size complexity can be achieved by a CRT-based approach with $W = O(n^2)$ [2].
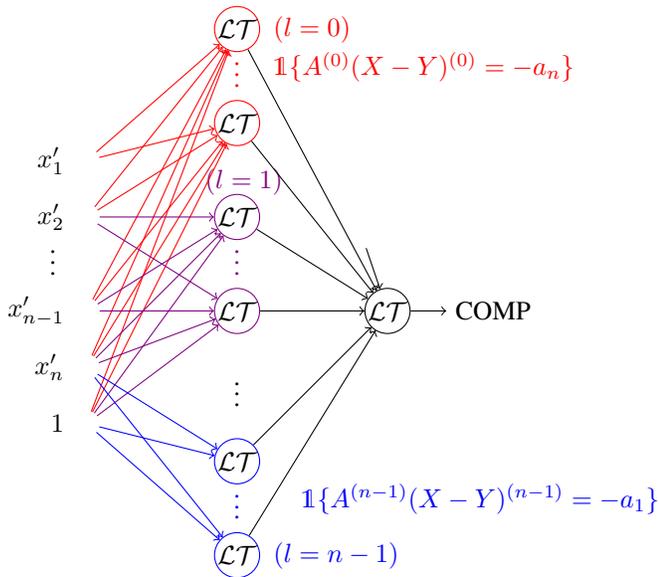
Fig. 5. A sketch of the COMP(X, Y) construction where $x_i' = x_i - y_i$ using linear threshold gates. Each color specifies an l value in the construction. If $X < Y$, all the rm many gates in at least one of the colors will give all 1s at the output. Otherwise, all the rm many gates in a color will give at most $(m - 1)$ many 1s at the output.

V. Conclusion

We explicitly constructed a rate-efficient constant weight EQ matrix for a previously existential result. Namely, we proved that the CRT matrix is not optimal in weight size to compute the EQ function and obtained the optimal EQ function constructions. For the COMP function, the weight size complexity is improved by a linear factor, using RMDS_q matrices and their existence. An open problem is whether similar algebraic constructions can be found for general threshold functions so that these ideas can be developed into a weight quantization technique for neural networks.

Acknowledgement

References

[1] N. Alon and V. H. Vũ, "Anti-hadamard matrices, coin weighing, threshold gates, and indecomposable hypergraphs," *Journal of Combinatorial Theory, Series A*, vol. 79, no. 1, pp. 133–160, 1997.

[2] K. Amano and A. Maruoka, "On the complexity of depth-2 circuits with threshold gates," in *International Symposium on Mathematical Foundations of Computer Science*, Springer, 2005, pp. 107–118.

[3] L. Babai, K. A. Hansen, V. V. Podolskii, and X. Sun, "Weights of exact threshold functions," in *International Symposium on Mathematical Foundations of Computer Science*, Springer, 2010, pp. 66–77.

[4] J. Beck, "Siegel's lemma is sharp," in *A journey through discrete mathematics*, Springer, 2017, pp. 165–206.

[5] T. Bohman, "A construction for sets of integers with distinct subset sums," *the electronic journal of combinatorics*, R3–R3, 1998.

[6] S. Costa, S. Della Fiore, and M. Dalai, "Variations on the Erdős Distinct-sums Problem," *arXiv preprint arXiv:2107.07885*, 2021.

[7] Q. Dubroff, J. Fox, and M. W. Xu, "A Note on the Erdős Distinct Subset Sums Problem," *SIAM Journal on Discrete Mathematics*, vol. 35, no. 1, pp. 322–324, 2021.

[8] M. Goldmann and M. Karpinski, "Simulating threshold circuits by majority circuits," in *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing*, 1993, pp. 551–560.

[9] R. Guy, *Unsolved problems in number theory*. Springer Science & Business Media, 2004, vol. 1.

[10] G. Halász, "Estimates for the concentration function of combinatorial number theory and probability," *Periodica Mathematica Hungarica*, vol. 8, no. 3-4, pp. 197–211, 1977.

[11] J. Håstad, "On the size of weights for threshold gates," *SIAM Journal on Discrete Mathematics*, vol. 7, no. 3, pp. 484–492, 1994.

[12] T. Hofmeister, "A note on the simulation of exponential threshold weights," in *International Computing and Combinatorics Conference*, Springer, 1996, pp. 136–141.

[13] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio, "Binarized neural networks," *Advances in neural information processing systems*, vol. 29, pp. 4107–4115, 2016.

[14] B. Jacob, S. Kligys, B. Chen, M. Zhu, M. Tang, A. Howard, H. Adam, and D. Kalenichenko, "Quantization and training of neural networks for efficient integer-arithmetic-only inference," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 2704–2713.

[15] S. R. Karingula and S. Lovett, "Singularity of random integer matrices with large entries," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2021)*, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.

[16] K. M. Kilic and J. Bruck, "Neural network computations with DOMINATION functions," in *2021 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2021, pp. 1029–1034.

[17] F. Li, B. Zhang, and B. Liu, "Ternary weight networks," *arXiv preprint arXiv:1605.04711*, 2016.

[18] D. Mandelbaum, "Error correction in residue arithmetic," *IEEE Transactions on Computers*, vol. 100, no. 6, pp. 538–545, 1972.

[19] S. Muroga, *Threshold logic and its applications*. 1971.

[20] V. P. Roychowdhury, A. Orlitsky, and K.-Y. Siu, "Lower bounds on threshold and related circuits via communication complexity," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 467–474, 1994.

[21] M. Rudelson and R. Vershynin, "The Littlewood–Offord problem and invertibility of random matrices," *Advances in Mathematics*, vol. 218, no. 2, pp. 600–633, 2008.

[22] C. L. Siegel, "Über einige anwendungen diophantischer approximationen," in *On Some Applications of Diophantine Approximations*, Springer, 2014, pp. 81–138.

[23] K.-Y. Siu and J. Bruck, "On the power of threshold circuits with small weights," *SIAM Journal on Discrete Mathematics*, vol. 4, no. 3, pp. 423–435, 1991.

[24] T. F. Tay and C.-H. Chang, "A non-iterative multiple residue digit error detection and correction algorithm in RRNS," *IEEE transactions on computers*, vol. 65, no. 2, pp. 396–408, 2015.

[25] G. Vardi and O. Shamir, "Neural networks with small weights and depth-separation barriers," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., vol. 33, Curran Associates, Inc., 2020, pp. 19 433–19 442. [Online]. Available: https : / / proceedings . neurips . cc / paper / 2020 / file / e1fe6165cad3f7f3f57d409f78e4415f-Paper.pdf.

The Decoding Algorithm: Suppose that $A_k x = z$ is given where A_k is an $m \times n$ EQ matrix given in Theorem 1 starting with $A_0 = \begin{bmatrix} 1 \end{bmatrix}$. One can obtain a linear time decoding algorithm in n by exploiting the recursive structure of the construction. Let us partition x and z in the way given in (13). It is clear that $x^{(3)} = [z^{(1)} + z^{(2)}]_2$. Also, after computing $x^{(3)}$, we find that $A_{k-1} x^{(1)} = (z^{(1)} + z^{(2)} - x^{(3)})/2$ and $A_{k-1} x^{(2)} = (z^{(1)} - z^{(2)} - x^{(3)})/2$. These operations can be done in $O(m_{k-1})$ time complexity. Let $T(m)$ denote the time to decode $z \in \mathbb{Z}^m$. Then, $T(m) = 2T(m/2) + O(m)$ and by the Master Theorem, $T(m) = O(m \log m) = O(n)$.

An Example for the Decoding: Consider the following system Ax = b:

$$
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\
1 & -1 & 0 & 1 & -1 & 0 & 0 & 1 \\
1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 \\
1 & -1 & 0 & -1 & 1 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8
\end{bmatrix}
=
\begin{bmatrix}
4 \\ -2 \\ -1 \\ 0
\end{bmatrix}
$$

Here is the first step of the algorithm:

$$
a = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} x_7 \\ x_8 \end{bmatrix} = [a+b]_2 = \begin{bmatrix} [3]_2 \\ [-2]_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}
$$

Then, we obtain the two following systems:

$$
\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{a + b - \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}
$$

$$
\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_4 \\ x_5 \\ x_6 \end{bmatrix} = \frac{a - b - \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{2} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}
$$

For the first system, we find that $x_3 = [1 - 1]_2 = 0$. Then,

$$
x_1 = \frac{1 - 1 - 0}{2} = 0 \tag{37}
$$

$$
x_2 = \frac{1 + 1 - 0}{2} = 1 \tag{38}
$$

For the second system, we similarly find that $x_6 = [2 - 1]_2 = 1$. Then,

$$
x_4 = \frac{2 - 1 - 1}{2} = 0 \tag{39}
$$

$$
x_5 = \frac{2 + 1 - 1}{2} = 1 \tag{40}
$$

Thus, we have $x = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}^T$.

A Modular Arithmetical Property of the CRT Matrix: Consider the following equation with powers of two in 8 variables and the corresponding 4×8 CRT matrix, say A.

$$
w_b^T x = \sum_{i=1}^{8} 2^{i-1} x_i = 0 \tag{41}
$$

$$
\begin{bmatrix}
[2^0]_3 & [2^1]_3 & [2^2]_3 & [2^4]_3 & [2^5]_3 & [2^6]_3 & [2^7]_3 \\
[2^0]_5 & [2^1]_5 & [2^2]_5 & [2^4]_5 & [2^5]_5 & [2^6]_5 & [2^7]_5 \\
[2^0]_7 & [2^1]_7 & [2^2]_7 & [2^4]_7 & [2^5]_7 & [2^6]_7 & [2^7]_7 \\
[2^0]_{11} & [2^1]_{11} & [2^2]_{11} & [2^4]_{11} & [2^5]_{11} & [2^6]_{11} & [2^7]_{11}
\end{bmatrix}
\tag{42}
$$

$$
= \begin{bmatrix}
1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 \\
1 & 2 & 4 & 3 & 1 & 2 & 4 & 3 \\
1 & 2 & 4 & 1 & 2 & 4 & 1 & 2 \\
1 & 2 & 4 & 8 & 5 & 10 & 9 & 7
\end{bmatrix}_{4 \times 8}
\tag{43}
$$

For the prime p_i, the elements in the row i are congruent to the elements of w_b. The ith element of the Ax vector should be divisible by p_i whenever $w_b^T x = 0$. For instance, one can pick $x = \begin{bmatrix} 2 & 1 & 1 & 3 & 0 & 1 & -1 & 0 \end{bmatrix}^T$ as a solution for $w_b^T x = 0$ and we see that $Ax = \begin{bmatrix} 12 & 15 & 14 & 33 \end{bmatrix}^T = \begin{bmatrix} 4 \cdot 3 & 3 \cdot 5 & 2 \cdot 7 & 3 \cdot 11 \end{bmatrix}^T$. This property is essential in the construction of small weight depth-2 circuits for arbitrary threshold functions while the RMDS$_q$ matrices do not behave in this manner necessarily.