

Weights of Exact Threshold Functions

László Babai¹, Kristoffer Arnsfelt Hansen^{*2},
Vladimir V. Podolskii^{**3}, and Xiaoming Sun^{***4}

¹ The University of Chicago
laci@cs.uchicago.edu

² Aarhus University
arnsfelt@cs.au.dk

³ Steklov Mathematical Institute
podolskii@mi.ras.ru

⁴ ITCS, Tsinghua University
xiaomings@tsinghua.edu.cn

Abstract. We consider Boolean exact threshold functions defined by linear equations, and in general degree d polynomials. We give upper and lower bounds on the maximum magnitude (absolute value) of the coefficients required to represent such functions. These bounds are very close. In the linear case in particular they are almost matching. This quantity is the same as the maximum magnitude of integer coefficients of linear equations required to express every possible intersection of a hyperplane in \mathbf{R}^n and the Boolean cube $\{0, 1\}^n$, or in the general case intersections of hypersurfaces of degree d in \mathbf{R}^n and the Boolean cube $\{0, 1\}^n$. In the process we construct new families of ill-conditioned matrices. We further stratify the problem (in the linear case) in terms of the dimension k of the affine subspace spanned by the solutions, and give upper and lower bounds in this case as well. Our bounds here in terms of k leave a substantial gap, a challenge for future work.

1 Introduction

A (linear) exact threshold function is a Boolean function that decides whether a real valued linear equality in its Boolean inputs holds. The related class of (linear) threshold functions consist of those Boolean functions that decide whether a real valued linear inequality in their Boolean inputs holds. To be more precise, an exact threshold function on n Boolean inputs x_1, \dots, x_n is a Boolean function that decides whether $w_1x_1 + \dots + w_nx_n = t$, where w_1, \dots, w_n are real valued

* Work done at Aarhus University supported by a postdoc fellowship from the Carlsberg Foundation and at The University of Chicago, supported by a Villum Kann Rasmussen postdoc fellowship.

** Partially supported by grant 09-01-00709-a from Russian Foundation for Basic Research Fund.

*** Supported by the National Natural Science Foundation of China Grant 60553001, 60603005 and 60621062 and the National Basic Research Program of China Grant 2007CB807900, 2007CB807901.

weights, and t is a real valued *threshold*. A threshold function on n Boolean inputs x_1, \dots, x_n is then a Boolean function that decides whether $w_1x_1 + \dots + w_nx_n \geq t$.

Threshold functions have been studied extensively in many areas of computer science (cf. [11, 14, 13]). Less attention has been given to exact threshold functions, but they have been considered in Boolean circuit complexity [4, 6–9] and in structural complexity theory [1]. We believe that studying exact threshold functions in itself is natural and interesting. However, an important reason for this is that such a study may bring additional insight to the study of threshold functions. For the results of [4, 6–8] proved for exact threshold functions, it is not currently known whether they also hold for threshold functions. A long standing open question for threshold functions is to prove good lower bounds for depth two circuits. Recent work [9] shows that circuit classes defined using exact threshold functions seamlessly interleave in the usual hierarchy of threshold circuit classes. In particular the class of depth two exact threshold circuits is shown to be a subclass of depth two threshold circuits. For this class no good lower bounds are known as well.

One can readily extend the notions of exact threshold functions as well as threshold functions to higher degree. A polynomial exact threshold function of degree d is a Boolean function that decides whether a real valued polynomial of degree d vanishes when evaluated on the Boolean input. Polynomial threshold functions are defined as an analogous generalization of threshold functions.

In this work we are interested in exact threshold functions from the fundamental perspective of representations of Boolean functions. More precisely, we are interested in the magnitude (absolute value) of integer weights needed to represent any possible exact threshold function. It is not hard to see that without loss of generality one may assume that the real valued weights and threshold defining an exact threshold function are in fact integers (as is also the case with threshold functions), thereby making the question we study well-defined.

The analogous question of the magnitude of weights required for threshold function has a long history of research. An upper bound on the magnitude of integer weights required to represent any threshold function was obtained by Muroga, Toda and Takasu [12] (cf. [13]). They showed that weights of magnitude $\leq (n+1)^{(n+1)/2}/2^n$ are sufficient. Also, several examples of explicit functions that require weights of magnitude $2^{\Omega(n)}$ are known [13, 14]. The existence of such functions may also be established by a counting argument, since there are at least $2^{n(n-1)/2}$ threshold functions on n variables [18, 17].

An almost optimal lower bound was obtained by Håstad [10]. Let n be a power of 2. Then Håstad constructed a threshold function requiring an integer weight of magnitude at least $(1/n)e^{-4n^\beta} n^{n/2}/2^n$, where $\beta = \log(3/2)$. Thus in the case when n is a power of 2 the upper bound and this lower bound differ only by a subexponential factor. Generalizing this work, for any constant d , [16] constructed a polynomial threshold function of degree d that requires integer weights of magnitude $n^{\Omega(n^d)}$.

Alon and Vü [2], building on the techniques of Håstad, gave a new construction of ill-conditioned matrices. For a non-singular $n \times n$ matrix A , let

$B = A^{-1} = (b_{ij})$ and define $\chi(A) = \max_{i,j} |b_{ij}|$. Define further $\chi_1(n)$ as the maximum of $\chi(A)$ over all non-singular $n \times n$ 0/1 matrices A . Define $\chi_2(n)$ to be the analogous quantity where $(-1, 1)$ matrices are considered instead.

With these definitions, Alon and Vü provide for every n an explicit $n \times n$ $(0, 1)$ matrix A_1 and an explicit $n \times n$ $(-1, 1)$ matrix A_2 such that $\chi(A_i) \geq n^{n/2}/2^{n(2-o(1))}$ for $i = 1, 2$. When n is a power of 2 these lower bounds may be improved to $n^{n/2}/2^{n(1-o(1))}$. Upper bounds for $\chi_i(n)$ are derived from the Hadamard inequality.

Theorem 1 (Alon and Vü). $n^{n/2}/2^{n(2-o(1))} \leq \chi_i(n)$ for $i = 1, 2$.
 $\chi_1(n) \leq n^{n/2}/2^{n-1}$, and $\chi_2(n) \leq (n-1)^{n-1}/2^{n-1}$.

Alon and Vü are able to apply their construction of ill-conditioned matrices to answer questions about flat simplices, weights of threshold functions, coin weighing, and indecomposable hypergraphs. In particular, they construct a threshold function on n variables that requires a weight of magnitude $\geq n^{n/2}/2^{n(2-o(1))}$.

Let $\maxw^T(n)$ denote the minimal W such that every possible threshold function on n variables can be realized using integer weights of magnitude $\leq W$. We summarize the above discussion in the following theorem.

Theorem 2 (Muroga et al.; Håstad; Alon and Vü).
 $n^{n/2}/2^{n(2-o(1))} \leq \maxw^T(n) \leq (n+1)^{n+1}/2^n$.

We are now in position to state our first theorem. We define $\maxw^E(n)$ for exact threshold functions as the analogous quantity of $\maxw^T(n)$. For this quantity we obtain the following upper and lower bounds.

Theorem 3. $n^{n/2}/2^{n(2-o(1))} \leq \maxw^E(n) \leq n^{n/2+1}$.

As is evident from Theorems 1, 2 and 3, the quantities $\chi_i(n)$, $\maxw^T(n)$ and $\maxw^E(n)$ are very close, in fact they are equal up to an exponential factor. Furthermore, when n is a power of 2, the bounds for $\chi_i(n)$ and $\maxw^T(n)$ differ only by a subexponential factor. We do not know if the same holds for $\maxw^E(n)$.

While some of our methods are related to the methods employed in the study of threshold functions, we do not see an explicit relationship between the quantities $\maxw^T(n)$ and $\maxw^E(n)$. The proofs of Theorem 2 and Theorem 3 in fact show that $\chi_2(n) \leq \maxw^T(n)$ and $\chi_1(n-1) \leq \maxw^E(n)$, but we do not know whether it is possible to turn a threshold function that requires certain magnitude into an exact threshold function requiring a similar magnitude or conversely.

A property that seems to be unique to (linear) exact threshold functions is that we can speak of the *dimension* of such functions. For a given exact threshold function f defined by a linear equation $w_1x_1 + \dots + w_nx_n = t$, consider the real affine space V generated by the points of the Boolean cube that satisfy the equation. We can then define the dimension of f to be the dimension of V .

Let $\maxw^E(n, k)$ denote the minimal W such that every exact threshold function of dimension k of n Boolean variables can be realized using integer weights of magnitude $\leq W$. Our second result gives the following for this quantity.

Theorem 4. For all n and all $1 \leq k \leq n$, $(\lfloor \frac{n}{k} \rfloor^k - 1)/2k \leq \max w_d^E(n, k) \leq n^{2^k}$.

For polynomial exact threshold functions, let $\max w_d^E(n)$ denote the magnitude of weights required to represent every possible exact degree d polynomial threshold function on n variables. Our final result is the following generalized bounds.

Theorem 5. $n^{\frac{1}{2}n^d}/2^{2n^d+o(n^d)+d} \leq \max w_d^E(2dn) \leq n^{\frac{dn^d}{2}+d}$.

This result is analogous to results for threshold functions, see [16]. For the lower bounds in this theorem we prove a specific generalization of Theorem 1.

The remainder of the paper is organized as follows. In Section 2 we state precisely the definitions we use and present some simple observations. In Section 3.1 we provide an example of a function of an exact threshold function that requires exponential magnitude of weights by an elementary argument. We prove Theorem 3 in Section 3.2. The proof of Theorem 4 is given in Sections 3.3 and 3.4 respectively. Theorem 5 is proved in Section 3.5. We conclude with open problems in Section 4.

2 Preliminaries

We consider here a Boolean function f to be a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$. We say that a Boolean function f on n variables is an exact threshold function if there exist real numbers w_1, \dots, w_n (the weights) and a real number t (the threshold value), such that $f(x) = 1$ if and only if $\sum_{i=1}^n w_i x_i = t$, for all $x \in \{0, 1\}^n$. We may say that the list of weight w_1, \dots, w_n and the threshold t as well as the expression $w_1 x_1 + \dots + w_n x_n = t$ are a *realization* of the function f . Similarly, a Boolean function f on n variables is a threshold function if there exist real numbers w_1, \dots, w_n and a real number t , such that $f(x) = 1$ if and only if $\sum_{i=1}^n w_i x_i \geq t$ for any $x \in \{0, 1\}^n$.

One may observe that without loss of generality it can be assumed that the real valued weights as well as the real valued threshold are in fact integers. Often when considering threshold functions one considers the Boolean cube $\{-1, 1\}^n$ instead of $\{0, 1\}^n$ as we do in this work. This is of no consequence to the possible weights of exact threshold and threshold functions. This is elaborated on in Appendix A. Furthermore, since, in this work we are only interested in the weights of exact threshold functions, we may without loss of generality restrict our attention to exact threshold functions f for which $f(0, \dots, 0) = 1$. Again, this is elaborated on in Appendix A.

Identifying a Boolean function f with the subset $f^{-1}(1)$, an exact threshold function corresponds to the intersection of a hyperplane in \mathbf{R}^n with the Boolean n -cube $\{0, 1\}^n$, and in the higher degree case an intersection with a degree d hypersurface. In the linear case, from the discussion above, for studying weights of exact threshold functions we may restrict our attention to affine spaces that are in fact subspaces of \mathbf{R}^n and are spanned by vectors of the Boolean cube $\{0, 1\}^n$. Because of this fact we will mainly phrase our theorems in vector space terminology.

3 Weights of an Exact Threshold Function

3.1 Example

An interesting example of an exact threshold function is the (sequence) equality function. Define the equality function EQ on $2n$ variables x_1, \dots, x_n and y_1, \dots, y_n by $\text{EQ}(x, y) = 1$ if and only if $x_i = y_i$ for all i .

It turns out that the set of weights that realize this exact threshold function corresponds *precisely* to solutions to the well known problem of finding n positive integers $a_1 < \dots < a_n$ such that all sums of the form $\sum_{i \in I} a_i$ are distinct. We explain this correspondence in Appendix B. Using this insight it is easy to see that weights of exponential magnitude are needed to realize the EQ function.

3.2 Upper and Lower Bounds for the Linear Case

Before proving the upper bound, we state three lemmas. Their proofs may be found in Appendices C, D and E.

Lemma 1 (Faddeev and Sominskii [5]). *Let A be a $n \times n$ matrix with all entries 0 or 1. Then $|\det(A)| \leq (n+1)^{\frac{n+1}{2}}/2^n$.*

Lemma 2. *Let $v_1, \dots, v_{n-1} \in \{0, 1\}^n$ be linearly independent. Then there exist integers w_1, \dots, w_n such that the equation $w_1x_1 + \dots + w_nx_n = 0$, defines the linear subspace $\text{span}(\{v_1, \dots, v_{n-1}\})$, and satisfy $|w_i| \leq n^{\frac{n}{2}}/2^{n-1}$.*

Lemma 3. *Let V be a vector space and let $k = n - \dim(\text{span}(V \cap \{0, 1\}^n))$. Then there exist vector spaces V_1, \dots, V_k such that $\dim(\text{span}(V_i \cap \{0, 1\}^n)) = n - 1$ for all i and $V \cap \{0, 1\}^n = \left(\bigcap_{i=1}^k V_i\right) \cap \{0, 1\}^n$.*

Theorem 6. *Let V be a vector space in \mathbf{R}^n . Then there exist integers w_1, \dots, w_n such that for all $x \in \{0, 1\}^n$ we have $x \in V$ if and only if $w_1x_1 + \dots + w_nx_n = 0$ and furthermore satisfy $|w_i| \leq n^{\frac{n}{2}+1}$ for all i .*

Thus every exact threshold function f on n variables can be realized using integer weights of absolute value at most $n^{\frac{n}{2}+1}$ as well.

Proof. Let $k = n - \dim(\text{span}(V \cap \{0, 1\}^n))$. Then by Lemma 3 there exist vector spaces V_1, \dots, V_k of dimension $n - 1$ spanned by vectors from $\{0, 1\}^n$ such that $V \cap \{0, 1\}^n = \left(\bigcap_{i=1}^k V_i\right) \cap \{0, 1\}^n$. For each V_i , by Lemma 2 there exist integers w_{ij} such that for all $x \in \{0, 1\}^n$ we have $x \in V_i$ if and only if the equation $w_{i1}x_1 + \dots + w_{in}x_n = 0$ is satisfied, and furthermore $|w_{ij}| \leq n^{\frac{n}{2}}/2^{n-1}$. We conclude the proof by the probabilistic method.

For $i = 1, \dots, k$, pick $c_i \in \{-2^{n-1}, \dots, 2^{n-1}\}$ uniformly at random, and consider the combined equation $\sum_{i=1}^k \sum_{j=1}^n c_i w_{ij} x_j = 0$. Clearly, every $x \in V \cap \{0, 1\}^n$ must satisfy this equation, since every such x satisfies the individual equations for all i .

Now consider $x \in \{0, 1\} \setminus V$. There must be some i such that $x \notin V_i$, and for this i the chosen x does not satisfy the corresponding equation. This implies

that the chosen x satisfies the combined equation with probability $< 2^{-n}$. Since $|\{0, 1\}^n \setminus V| \leq 2^n$, the probability that some $x \in \{0, 1\}^n \setminus V$ satisfies the combined equation is < 1 , and thus there is a fixed choice \hat{c} for c with this property. Hence, $w_j = \sum_{i=1}^k \hat{c}_i w_{ij}$ satisfies the requirements, since each $|\hat{c}_i w_{ij}| \leq n^{\frac{n}{2}}$. \square

Ziegler [19] using the results of Theorem 1 gave a lower bound on the maximal coefficient of an inequality defining a facet of a full-dimensional $(0, 1)$ polytope in \mathbf{R}^n . Since the polytope is of full dimension the hyperplane given by the facet is uniquely determined by points from $\{0, 1\}^n$ and hence corresponds to a unique exact threshold function, and the lower bound applies to our setting also. We state the lower bound below, and additionally point out that the construction provides a lower bound on the magnitude of *all* the coefficients.

Theorem 7. *For any n , there exists an exact threshold function f on n variables such that any realization of f requires an integer weight of magnitude $n^{\frac{n}{2}}/2^{n(2-o(1))}$.*

The proof is given in Appendix F

Observation 1. Alon and Vũ show when $n - 1$ is a power of 2, that the inverse of the $(n - 1) \times (n - 1)$ matrix A they construct actually has a column (in fact many columns) where *all* entries are of magnitude $n^{\frac{n}{2}}/2^{\Theta(n)}$. This means that in the construction above, when $n - 1$ is a power of 2, one can obtain that all the first $n - 1$ coefficients are of this magnitude.

In other words, for every n , not only is it such that one may be required to use $\Omega(n \log n)$ bits to store the largest weight, but to store all the weights one may be required to use $\Omega(n^2 \log n)$ bits.

3.3 Small Dimension Upper Bound

Theorem 8. *Let V be a vector space in \mathbf{R}^n and let $k = \dim(\text{span}(V \cap \{0, 1\}^n))$. Then there exist integers w_1, \dots, w_n such that for all $x \in \{0, 1\}^n$ we have $x \in V$ if and only if $w_1 x_1 + \dots + w_n x_n = 0$ and furthermore satisfy $|w_i| \leq n^{2^k}$ for all i . Thus every exact threshold function f on n variables of dimension at most k can be realized using integer weights of absolute value at most n^{2^k} as well.*

Proof. Let $v_1, \dots, v_k \in \{0, 1\}^n$ be a basis of $\text{span}(V \cap \{0, 1\}^n)$. For $\alpha \in \{0, 1\}^k$ define the set $S_\alpha = \{i \in \{1, \dots, n\} \mid \forall j \in \{1, \dots, k\} : (v_j)_i = \alpha_j\}$. Number the nonempty such sets S_1, \dots, S_K for $K \leq 2^k$. For $i \in \{1, \dots, K\}$, let $n_i = |S_i|$, and assume $S_i = \{a_{i1}, \dots, a_{in_i}\}$. Further, define $\chi_i \in \{0, 1\}^n$ to be the characteristic vector of the set S_i .

Now, consider the vector space W in \mathbf{R}^K defined by $y \in W$ if and only if $\sum_{i=1}^K y_i \chi_i \in V$. Note that, if $\sum_{i=1}^K y_i \chi_i \in \{0, 1\}^n$, we must have $y \in \{0, 1\}^K$ by construction. By Theorem 6 we have integer weights $\hat{w}_1, \dots, \hat{w}_K$ such that for all $y \in \{0, 1\}^K$ we have $y \in W$ if and only if $\hat{w}_1 y_1 + \dots + \hat{w}_K y_K = 0$ and furthermore satisfy $|\hat{w}_i| \leq K^{\frac{K}{2}+1}$. Thus, we also have for all $y \in \{0, 1\}^K$ it holds that $\sum_{i=1}^K y_i \chi_i \in V$ if and only if $\hat{w}_1 y_1 + \dots + \hat{w}_K y_K = 0$.

We now define integer weights w_1, \dots, w_n as follows. Let $N = \prod_{l=1}^K n_l$. If $n_i = 1$ we simply define $w_{a_{i1}} = N\hat{w}_i$. Otherwise we give the first element of each set S_i weight as $w_{a_{i1}} = -(n_i - 1) \prod_{l=1}^{i-1} n_l + N\hat{w}_i$, and the remaining elements are given weights as $w_{a_{ij}} = \prod_{l=1}^{i-1} n_l$, for $j = 2, \dots, n_i$. By construction we obtain the property that if $w_1 x_1 + \dots + w_n x_n = 0$ for $x \in \{0, 1\}^n$ we must have that for all i we have that all coordinates $x_{a_{ij}}$ have the same value, 0 or 1. Thus x must be of the form $\sum_{i=1}^K y_i \chi_i \in W$ for $y \in \{0, 1\}^K$. The converse trivially holds. Now, finally note that for all i we have $|w_i| \leq K^{\frac{K}{2}+1} \prod_{j=1}^K n_j \leq K^{\frac{K}{2}+1} \left(\frac{n}{K}\right)^K = \frac{n^K}{K^{\frac{K}{2}-1}} \leq n^K \leq n^{2^k}$. \square

3.4 Small Dimension Lower Bound

Suppose $k \leq n/2$, let $d = \lfloor \frac{n}{k} \rfloor$, and define a k dimensional vector space V by

$$V = \text{span}\left\{ \overbrace{1 \dots 1}^d \overbrace{0 \dots 0}^{n-d}, \overbrace{0 \dots 0}^d \overbrace{1 \dots 1}^d \overbrace{0 \dots 0}^{n-2d}, \dots, \overbrace{0 \dots 0}^{(k-1)d} \overbrace{1 \dots 1}^d \overbrace{0 \dots 0}^{n-kd} \right\}.$$

Theorem 9. *Suppose w_1, \dots, w_n are integers satisfying $x \in V$ if and only if $\sum_{i=1}^n w_i x_i = 0$ for all $x \in \{0, 1\}^n$. Then we must have $\max_i |w_i| \geq \frac{d^k - 1}{2k} \sim \frac{n^k}{2k^{k+1}}$. Thus there exist an exact threshold function on n variables of dimension k that requires an integer weight of absolute value at least $(d^k - 1)/2k$ as well.*

Proof. First we relabel the weights, w_1, \dots, w_{kd} by $w_{1,1}, \dots, w_{1,d}; w_{2,1}, \dots, w_{2,d}; \dots; w_{k,1}, \dots, w_{k,d}$, where $w_{i,j} = w_{(i-1)d+j}$ ($i \in [k], j \in [d]$). For each $i \in [k]$ define the (multi)-set S_i by $S_i = \{w_{i,1}, w_{i,2}, \dots, w_{i,d}\}$. For a (multi)-set $S \subset \mathbf{Z}$ define $\text{sum}(S) = \sum_{y \in S} y$. By assumption and by the definition of V we have $\text{sum}(S_i) = 0$, for all i . We claim that if $\max_i |w_i| < \frac{d^k - 1}{2k}$, then there exist subsets $\tilde{S}_i \subseteq S_i$ for each $i \in [k]$, such that $\sum_{i=1}^k \text{sum}(\tilde{S}_i) = 0$, and at least one \tilde{S}_i satisfies $\tilde{S}_i \neq \emptyset, \tilde{S}_i \neq S_i$. In other words there would exist $x \in \{0, 1\}^n$ satisfying $\sum_{i=1}^n w_i x_i = 0$ and $x \notin V$, leading to a contradiction. Hence we can conclude that $\max_i |w_i| \geq \frac{d^k - 1}{2k}$. We next prove this claim, thereby completing the proof of the theorem.

Let $M = \max_i |w_i|$. By assumption we have $M < \frac{d^k - 1}{2k}$. Since $\text{sum}(S_i) = 0$, i.e. $w_{i,1} + w_{i,2} + \dots + w_{i,d} = 0$, we can arrange $\{w_{i,1}, \dots, w_{i,d}\}$ in an order $\{\tilde{w}_{i,1}, \dots, \tilde{w}_{i,d}\}$ such that: (1) $\tilde{w}_{i,1} \geq 0$; (2) for each $2 \leq j \leq d$, if $\sum_{l=1}^{j-1} \tilde{w}_{i,l} \geq 0$, then $\tilde{w}_{i,j} \leq 0$, otherwise $\tilde{w}_{i,j} > 0$. It is easy to see that for any $i \in [k], j \in [d]$ we have, $-M \leq \tilde{w}_{i,1} + \tilde{w}_{i,2} + \dots + \tilde{w}_{i,j} \leq M$.

Now, for each k -tuple $(l_1, \dots, l_k) \in [d]^k$, consider the double summation $\sum_{i=1}^k \sum_{j=1}^{l_i} \tilde{w}_{i,j}$. By the previous equation we have $-kM \leq \sum_{i=1}^k \sum_{j=1}^{l_i} \tilde{w}_{i,j} \leq kM$. But in total there are d^k different k -tuples (l_1, \dots, l_k) , and $2kM + 1 < 2k \cdot \frac{d^k - 1}{2k} + 1 = d^k$. Therefore there exist two k -tuples $(l_1, \dots, l_k) \neq (l'_1, \dots, l'_k)$, such that $\sum_{i=1}^k \sum_{j=1}^{l_i} \tilde{w}_{i,j} = \sum_{i=1}^k \sum_{j=1}^{l'_i} \tilde{w}_{i,j}$.

For each $i \in [k]$, we define a set \tilde{S}_i as follows. If $l_i \geq l'_i$, let $\tilde{S}_i = \{\tilde{w}_{i,l'_i+1}, \dots, \tilde{w}_{i,l_i}\}$. Otherwise, let $\tilde{S}_i = S_i \setminus \{\tilde{w}_{i,l_i+1}, \dots, \tilde{w}_{i,l'_i}\}$. Combining the previous equation with the fact that $\text{sum}(S_i) = 0$, we obtain $\sum_{i=1}^k \text{sum}(\tilde{S}_i) = 0$.

It is clear that $\tilde{S}_i \neq S_i$ for all i , and since $(l_1, \dots, l_k) \neq (l'_1, \dots, l'_k)$, there must exist i such that $\tilde{S}_i \neq \emptyset$. \square

3.5 Higher Degree Bounds

The upper bound for polynomial exact threshold functions follows easily from the upper bound for the linear case. We give the proof in Appendix G.

As to the lower bound we proceed similarly to the case of degree 1. We start by a specific generalization of Alon and Vü's lower bound on $\chi(A)$ [2]. This could be done by translating the analysis of [16] to matrix terminology analogous of Alon and Vü's translated analysis of Håstad's result [10]. But instead we shall prefer to present a technically simpler proof using matrix terminology based on the results of [2]. We note, however, that although this proof differs from the proof in [16], the underlying ideas are the same.

In what follows we use parameters n and d , where d denotes the degree of the polynomial and nd is the number of input variables. We denote input variables by x_{ij} for $i = 1, \dots, d$, $j = 1, \dots, n$ and suppose they range over $\{0, 1\}$. Let $\bar{x}^1 = (x_{1,1}, \dots, x_{1,n}), \dots, \bar{x}^d = (x_{d,1}, \dots, x_{d,n})$. Let $x = (\bar{x}^1, \dots, \bar{x}^d)$. We denote by M_d the set of monomials $\{x_{1,j_1} x_{2,j_2} \cdots x_{d,j_d} \mid j_1, \dots, j_d \in \{1, \dots, n\}\}$ (that is, we take exactly one variable from each \bar{x}^i). For the matrix A we denote by A_{ij} its submatrix obtained by deleting the i th row and the j th column.

To state our generalization we first need the following definition.

Definition 1. *A matrix $A \in \{0, 1\}^{m \times m}$ is called d -generable if we can label each column of A with a unique monomial from M_d and label each row of A by an assignment to the variables (an input) in such a way that each entry a_{ij} is exactly the value of the monomial corresponding to the column j when evaluated on the input corresponding to the row i .*

Now we can provide the generalization of Alon and Vü's lower bound.

Theorem 10. *For any d and any n there exist an (explicit) matrix $A^{(d)} \in \{0, 1\}^{n^d \times n^d}$ such that $A^{(d)}$ is d -generable, $|\det A_{1,n^d}^{(d)} / \det A^{(d)}| \geq n^{\frac{1}{2}n^d} / 2^{2n^d + o(n^d)}$ and $A^{(d)}$ has the form $\begin{bmatrix} e_1^T \\ B \end{bmatrix}$, where e_1^T is a row $(1, 0, \dots, 0)$ of length n^d and B is an $(n^d - 1) \times n^d$ matrix. (Note that the function hidden in o -notation above depends on n but does not depend on d .)*

Proof (sketch). The proof goes by induction on d . In the base case, $d = 1$, the matrix $A^{(1)}$ can be easily obtained from the matrix constructed in [2]. For inductive step we need the following notation.

For a matrix $A = \{a_{ij}\}_{i,j=1}^n$, let \bar{A} denote the matrix obtained by reflecting A in the vertical median, i.e., $\bar{A} = \{a_{i,n+1-j}\}_{i,j=1}^n$. Now we define $A^{(d)}$ by

$$A^{(d)} = \begin{bmatrix} A^{(d-1)} & 0 & 0 & \cdots \\ e_n^T & \bar{A}^{(d-1)} & 0 & \cdots \\ 0 & e_1^T & A^{(d-1)} & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}$$

This matrix has $n \times n$ blocks. In the diagonal blocks we have matrices $A^{(d-1)}$, $\bar{A}^{(d-1)}$, $A^{(d-1)}$, \dots . In the blocks right below the diagonal only the first row is nonzero and these rows are $e_n^T, e_1^T, e_n^T, \dots$. All other blocks consists exclusively of zeros.

It is now just a matter of checking that $A^{(d)}$ satisfies all required properties. To verify the claimed inequality, we repeatedly apply Lemma 2.3.1 from [2]. The details of the proof are given in the Appendix H. \square

Now we prove the main result of this section.

Theorem 11. $\max w_d^E(2dn) \geq n^{\frac{1}{2}n^d} / 2^{2n^d + o(n^d) + d}$.

Proof. We will now have variables $x_{ij}, y_{ij} \in \{0, 1\}$ for $i = 1, \dots, d$ and $j = 1, \dots, n$ (that is, we have twice the number of variables), and consider integer polynomials $p(x, y)$ of degree d . In fact it will be more convenient for us to consider polynomials $q(x - y, x + y)$ instead (recall that x and y are vectors of variables). It is easy to check that we can always convert a polynomial $p(x, y)$ into an equivalent polynomial $q(x - y, x + y)$ and vice versa. Moreover, if $q(x - y, x + y)$ has a large coefficient then the corresponding polynomial $p(x, y)$ will also have a large coefficient.

Observation 2. Suppose all coefficients of a degree- d integer polynomial $p(x, y)$ are of absolute value at most s . Let $q(u, v) = p((v + u)/2, (v - u)/2)$. Then $p(x, y) = q(x - y, x + y)$, and $2^d q(u, v)$ has integer coefficients of absolute value at most $2^d s$. \square

We will now construct a polynomial exact threshold function with the desired properties. We start with the matrix $A^{(d)}$ given by Theorem 10. First of all we switch the labeling of the matrix: in each monomial in the labeling of $A^{(d)}$ we substitute each variable x_{ij} by the expression $x_{ij} - y_{ij}$. Thus now our columns correspond to monomials in variables $x - y$. Let us denote this new set of monomials corresponding to columns of $A^{(d)}$ by M'_d . To complete the labeling

we must change the row labels correspondingly. This is possible since we can find values of x_{ij} and y_{ij} in such a way that the old value of x_{ij} is equal to the new value of $x_{ij} - y_{ij}$. Now we proceed with a proof.

Let us consider the matrix $B = [A^{(d)} \ e_1]$. That is, we add one column to the matrix $A^{(d)}$. Let $z \in \{0, 1\}$ be a new variable and let the new column correspond to the monomial $\theta = (x_{11} - y_{11})(x_{21} - y_{21}) \dots (x_{d1} - y_{d1})z$. We next need to extend the assignments labeling the rows to include values of z . For the first row we let $z = 1$ and for the others rows we let $z = 0$. Now it is easy to see that each entry of B is equal to the value of monomial corresponding to the column on the assignment corresponding to the row.

Let us now consider an exact polynomial threshold gate over the set of monomials $M'_d \cup \{\theta\}$ coefficient vector w of which is a nonzero integer solution of the system $Bw = 0$. Note that the dimension of the solution space of this system is 1 since $A^{(d)}$ is nonsingular and thus w is uniquely determined up to a multiplicative factor. We denote this polynomial by $p(x - y, z)$ and the corresponding exact threshold function by $f(x, y, z)$.

Now we will prove that any integer representation w' of the same exact threshold function over the set of all monomials over the variables $x - y$, $x + y$ and over the monomial θ must have a large coordinate. First, note that such representation should satisfy the system

$$[B \ B'] \begin{pmatrix} u \\ v \end{pmatrix} = 0, \quad (1)$$

where the columns B' corresponds to monomials in variables $x - y$ and $x + y$ which are not in M'_d and $\begin{pmatrix} u \\ v \end{pmatrix}$ is w' where u corresponds to the B -part and v corresponds to the B' -part. Entries of the matrix B' are as usual equal to the value of the monomial corresponding to the column on the input corresponding to the row. So, B' is $\{0, \pm 1\}$ -matrix.

Proposition 1. $Bu = 0$.

The proof of this proposition uses a symmetry argument and can be found in the Appendix I.

Now if we detach the last column of the matrix and move it to the right-hand side we will get system with the square matrix:

$$A^{(d)} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n^d} \end{pmatrix} = \begin{pmatrix} u_{n^d+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Note that u_{n^d+1} should be nonzero since otherwise all u_i are zero (recall that $A^{(d)}$ is nonsingular). Now by Cramer's rule we have that $|u_{n^d}| = \left| \frac{u_{n^d+1} \det A_{1, n^d}^{(d)}}{\det A^{(d)}} \right| \geq$

$\left| \frac{\det A_{1,n^d}^{(d)}}{\det A^{(d)}} \right|$. Now we get rid of z . Define the function $g(x, y) = f(x, y, 0)$. It is obviously a polynomial exact threshold function. Moreover any degree- d integer polynomial representing $g(x, y)$ can be transformed into a degree- d integer polynomial representing $f(x, y, z)$. Indeed, let the polynomial for g be given by the vector of coefficients u . Add to it an integer coefficient for the monomial θ in such a way that resulting vector u' satisfy the first row of the system (1). Note also that it automatically satisfies all other rows of this system (since they refer to inputs with $z = 0$). Thus u' is a solution of the system (1) and hence

$$|u_{n^d}| \geq \left| \frac{\det A_{1,n^d}^{(d)}}{\det A^{(d)}} \right| \geq \frac{n^{\frac{1}{2}n^d}}{2^{2n^d + o(n^d)}}.$$

Now we have proved the lower bound on the coefficient of an integer polynomial of degree at most d in variables $x - y$ and $x + y$ which represents $g(x, y)$. And by Lemma 2 we have desired lower bound for polynomials in variables x and y (here a factor 2^d appears in the denominator). \square

Remark 1. Note that the place where we needed a new variable z is Proposition 1. If we should try instead to write a system without the variable z and with nontrivial right-hand side we would not be able to perform the symmetry argument.

Remark 2. We can reprove lower bound on the $\max w_d^T(n)$ (which is a generalization of $\max w^T(n)$ to degree- d threshold functions) from [16] similarly generalizing proof of the lower bound on $\max w^T(n)$ from [2]. See Appendix J.

4 Conclusion

We have obtained upper and lower bounds for the magnitude of integer weights required to represent exact threshold functions, for linear exact threshold functions as well as polynomial exact threshold functions in general. For the linear case, we also gave bounds for the interesting special case of small dimension functions. In the small dimension case there seems to be ample room for further improvement of the bounds. In the other cases our bounds are very close, especially for the linear case, leaving little room for improvement. However our proof raises an interesting question: Is it possible that exact threshold functions on n variables of dimension less than $n - 1$ can require larger weights than those of dimension $n - 1$, or is the worse upper bound an artifact of our proof. For obtaining the lower bounds for polynomial exact threshold functions we constructed ill-conditioned matrices with special properties – these may have additional applications.

Another question arises from comparing results known for threshold functions and for exact threshold functions. Beigel [3] give a technique for showing that a simple linear threshold function require exponential weights even for representations of degree $n^{1/2-\epsilon}$. This technique was generalized in [15] to the case of higher degree polynomial threshold functions. Beigel's approach and its extensions does not seem to be applicable to exact threshold functions. Obtaining analogous results for exact threshold functions is therefore an open problem.

References

1. Agrawal, M., Arvind, V.: Geometric sets of low information content. *Theoretical Computer Science* 158(1–2), 193–219 (1996)
2. Alon, N., Vü, V.H.: Anti-hadamard matrices, coin weighing, threshold gates, and indecomposable hypergraphs. *Journal of Combinatorial Theory, Series A* 79(1), 133–160 (1997)
3. Beigel, R.: Perceptrons, PP, and the polynomial hierarchy. *Computational Complexity* 4(4), 339–349 (1994)
4. Beigel, R., Tarui, J., Toda, S.: On probabilistic ACC circuits with an exact-threshold output gate. In: *Proceedings of the Third International Symposium on Algorithms and Computation*. Lecture Notes in Computer Science, vol. 650, pp. 420–429. Springer (1992)
5. Faddeev, D.K., Sominskii, I.S.: *Problems in Higher Algebra*. W. H. Freeman (1965), (translated by J. L. Brenner)
6. Green, F.: A complex-number fourier technique for lower bounds on the mod- m degree. *Computational Complexity* 9(1), 16–38 (2000)
7. Hansen, K.A.: Computing symmetric boolean functions by circuits with few exact threshold gates. In: *Proceedings of the 13th Annual International Conference on Computing and Combinatorics*. Lecture Notes in Computer Science, vol. 4598, pp. 448–458. Springer (2007)
8. Hansen, K.A.: Depth reduction for circuits with a single layer of modular counting gates. In: *Proceedings of the Fourth International Computer Science Symposium in Russia*. Lecture Notes in Computer Science, vol. 5675, pp. 117–128. Springer (2009)
9. Hansen, K.A., Podolskii, V.V.: Exact threshold circuits. In: *Proceedings of the 25th Annual IEEE Conference on Computational Complexity*. IEEE (2010), to appear
10. Hästad, J.: On the size of weights for threshold gates. *SIAM Journal on Discrete Mathematics* 7(3), 484–492 (1994)
11. Hertz, J., Krogh, A., Palmer, R.G.: *Introduction to the Theory of Neural Computation*. Addison-Wesley Publishing Company (1991)
12. Muroga, S., Toda, I., Takasu, S.: Theory of majority decision elements. *Journal of the Franklin Institute* 271, 376–418 (1961)
13. Muroga, S.: *Threshold Logic and its Applications*. John Wiley & Sons, Inc. (1971)
14. Parberry, I.: *Circuit Complexity and Neural Networks*. MIT Press, Cambridge, MA (1994)
15. Podolskii, V.V.: A uniform lower bound on weights of perceptrons. In: *Proceedings of the 3rd International Computer Science Symposium in Russia*. Lecture Notes in Computer Science, vol. 5010, pp. 261–272. Springer (2008)
16. Podolskii, V.V.: Perceptrons of large weight. *Problems of Information Transmission* 45(1), 46–53 (2009)
17. Smith, D.R.: Bounds on the number of threshold functions. *IEEE Transactions on Electronic Computers* EC–15(6), 368–369 (1966)
18. Yajima, S., Ibaraki, T.: A lower bound on the number of threshold functions. *IEEE Transactions on Electronic Computers* EC–14(6), 926–929 (1965)
19. Ziegler, G.M.: Lectures on 0/1-polytopes. In: Kalai, G., Ziegler, G.M. (eds.) *Polytopes - Combinatorics and Computation*, DMV Seminar, vol. 29, pp. 1–43. Birkhäuser (2000)

Appendix

Remark: Additional references used in the appendix are found at the end of the appendix.

A Elaboration on Preliminaries

One may observe that without loss of generality one may assume that the real valued weights as well as the real valued threshold are in fact integers. Often when considering threshold functions one chooses to consider the Boolean cube $\{-1, 1\}^n$ instead of $\{0, 1\}^n$ as we do in this work. This is of no consequence to the possible weights of exact threshold and threshold functions. Suppose $f(x_1, \dots, x_n)$ is such a function in variables $x_1, \dots, x_n \in \{0, 1\}$. Then substituting $(y_i + 1)/2$ for x_i for all i and multiplying the resulting expression by 2 yields an equivalent function in variables $y_1, \dots, y_n \in \{-1, 1\}^n$ with the exact same weights. Suppose on the other hand that $f(y_1, \dots, y_n)$ is a such a function in variables $y_1, \dots, y_n \in \{-1, 1\}^n$. Substitute $2x_i - 1$ for y_i and expand the new expression. If the new constant coefficient is even we may divide the expression by 2. If the constant coefficient is odd, we may in case of a threshold function divide by 2 and round the new threshold value up to the nearest integer. In case of an exact threshold function the expression can never be satisfied, making the function trivial. We may thus for example replace the new constant coefficient with a sufficiently large even integer and divide the expression by 2. In all cases we obtain an equivalent function in variables $x_1, \dots, x_n \in \{0, 1\}$ with the exact same weights.

Since, in this work we are only interested in the weights of exact threshold functions, we may without loss of generality restrict our attention to exact threshold functions f for which $f(0, \dots, 0) = 1$. This may be seen as follows. Suppose f is a non-constant exact threshold function where $f(0, \dots, 0) = 0$, and consider any realization of f . Since f is non-constant, let $\hat{x} \in \{0, 1\}^n$ be such that $f(\hat{x}) = 1$. Whenever $\hat{x}_i = 1$ substitute $1 - y_i$ for x_i and when $\hat{x}_i = 0$ substitute y_i for x_i . This yields a realization of a new exact threshold function g that satisfy $g(0, \dots, 0) = 1$. Now note that the sets of magnitudes of weights of the two realizations are exactly the same.

B The EQ exact threshold function

Proposition 2. *Suppose that EQ is realized as an exact threshold function by $\sum_{i=1}^n w_i x_i + \sum_{i=1}^n w'_i y_i = t$. Then $t = 0$ and $w_i = -w'_i$ for all i . Thus EQ is also realized by the expression $\sum_{i=1}^n |w_i|(x_i - y_i) = 0$.*

Proof. Clearly $t = 0$ since $\text{EQ}(0 \dots 0, 0 \dots 0) = 1$. Now, for a given i , let $x_i = y_i = 1$ and $x_j = y_j = 0$ for $j \neq i$. Since $\text{EQ}(x, y) = 1$ we also have $w_i = -w'_i$.

Now, given x and y , let $x'_i = x_i$ and $y'_i = y_i$ if $w_i > 0$ and let $x'_i = y_i$ and $y'_i = x_i$ otherwise. We then have $\sum_{i=1}^n |w_i|(x_i - y_i) = 0 \Leftrightarrow \sum_{i=1}^n w_i(x'_i - y'_i) = 0 \Leftrightarrow x' = y' \Leftrightarrow x = y$. \square

We thus have the following characterization of representations of EQ as an exact threshold function.

Corollary 1. *Suppose that EQ is realized by an exact threshold function given by $\sum_{i=1}^n w_i x_i + \sum_{i=1}^n w'_i y_i = t$. Then all sums of the form $\sum_{i \in I} |w_i|$ for $I \subseteq [n]$ must be distinct. Conversely, if w_1, \dots, w_n are positive real numbers such that all sums of the form $\sum_{i \in I} w_i$ for $I \subseteq [n]$ are distinct, then $\sum_{i=1}^n w_i(x_i - y_i) = 0$ defines EQ.*

Thus the problem of determining the magnitude of integers weights required for realizing the EQ function on $2n$ variables is precisely the well known problem of finding n positive integers $a_1 < \dots < a_n$ such that all sums of the form $\sum_{i \in I} a_i$ are distinct. Clearly we can obtain $a_n \leq 2^{n-1}$ and it is also easy to see that $a_n \geq \frac{2^n}{n}$ is necessary. Erdős conjectured that $a_n \geq c2^n$ for some constant c , but this remains unresolved. Erdős and Moser [24] proved that $a_n \geq \frac{2^n}{4\sqrt{n}}$. Conway and Guy [22] showed that $a_n < 0.23513 \cdot 2^n$. Since then both the lower and upper bound have been improved, albeit only by constant factors. The upper bound has been improved by Bohman [21] to $a_n < 0.22002 \cdot 2^n$, for sufficiently large n . The lower bound was improved to $a_n > 2^{-n} \binom{2n}{n} \sim \frac{2^n}{\sqrt{\pi n}}$ by Elkies [23], and this have been further improved asymptotically by a factor $\sqrt{\frac{3}{2}}$ by Aliev and a factor $\sqrt{2}$ by Elkies and Gleason [20].

C Proof of Lemma 1

Proof. Following Williamson [25], define the $(n+1) \times (n+1)$ matrix \hat{A} with all entries -1 or 1 by

$$\hat{A} = \begin{bmatrix} 1 & -\mathbf{1}^T \\ \mathbf{1} & 2A - J \end{bmatrix},$$

where J is the $n \times n$ matrix which have all entries 1. Adding the first column to the other columns and then expanding the determinant from the first row we see that

$$\det(\hat{A}) = 2^n \det(A).$$

Applying the Hadamard inequality to \hat{A} we obtain

$$\det(\hat{A})^2 \leq \prod_{i=1}^{n+1} \left(\sum_{j=1}^{n+1} (\hat{a}_{ij})^2 \right) \leq (n+1)^{n+1}.$$

Combining the two equations we obtain the result. \square

D Proof of Lemma 2

Proof. Let A be the $(n-1) \times n$ matrix, whose rows are the vectors v_1, \dots, v_{n-1} . Then w must be any nonzero solution to the linear system $Aw = 0$. Let A_i be the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i th column. Then a solution is given by $w_i = (-1)^{i+1} \det(A_i)$. Since the entries of the matrices A_i are 0 or 1 the result follows using Lemma 1. \square

E Proof of Lemma 3

Proof. Let \mathcal{W} be the collection of vector spaces W satisfying that $V \cap \{0, 1\}^n \subset W$ and $\dim(\text{span}(W \cap \{0, 1\}^n)) = n-1$. By dimension arguments it is sufficient to show that $V \cap \{0, 1\}^n = (\bigcap_{W \in \mathcal{W}} W) \cap \{0, 1\}^n$.

Let $x \in \{0, 1\} \setminus V$ be arbitrary. We will construct $W \in \mathcal{W}$ such that $x \notin W$, and this implies the above statement. This can be done by the following iterative procedure. For $i = 1, \dots, n-1$, since $\{0, 1\}^n \not\subseteq \text{span}(\{v_1, \dots, v_{i-1}, x\})$ we can pick $v_i \in \{0, 1\}^n \setminus \text{span}(\{v_1, \dots, v_{i-1}, x\})$, and we have $x \notin \text{span}(\{v_1, \dots, v_i\})$. After this procedure, simply let $W = \text{span}(\{v_1, \dots, v_{n-1}\})$. \square

F Proof by Ziegler

Proof. (Ziegler)

Let A be the $(n-1) \times (n-1)$ matrix given by Theorem 1. Assume without loss of generality that $\chi(A) = \frac{\det(A_{11})}{\det(A)} \geq \frac{n^{\frac{n}{2}}}{2^{n(2-o(1))}}$. Form the $(n-1) \times n$ matrix \hat{A} by

$$\hat{A} = [A \ e_1] \quad ,$$

As in Lemma 2 we get that weights w for the exact threshold function defined by the rows of \hat{A} must be a solution to the linear system $\hat{A}w = 0$. One such solution is given by $w_i = (-1)^{i+1} \det(\hat{A}_i)$, and any solution is a scalar multiple of this. Now, $w_1 = \det(A_{11})$ and $w_n = (-1)^{n+1} \det(A)$, and thus $|w_1/w_n| = \chi(A)$. Thus any integer solution \hat{w} must have $|\hat{w}_1| \geq \chi(A)$. \square

G Proof of Upper Bound of Theorem 5

Consider some polynomial exact threshold function of degree d on n variables. Let $p(x)$ be a corresponding integer degree- d polynomial. Substitute each monomial by a new fresh variable u_i . There are at most n^d monomials in p (since it is assumed to be multilinear) so we need at most n^d new variables $u = (u_1, \dots, u_{n^d})$. Thus we obtained a linear polynomial in the variables u_i . Denote this polynomial by $q(u)$, and consider the linear exact threshold function corresponding to q . By Theorem 6 we obtain a linear polynomial $q'(u)$ corresponding to this function with integer weights of absolute value at most $(n^d)^{\frac{n^d}{2}+1} = n^{\frac{dn^d}{2}+d}$. Now we simply substitute back in the polynomial $q'(u)$ all

variables u_i by the corresponding monomials in the variables x_i . Doing this we obtain polynomial $p'(x)$ with integer weights of absolute value at most $n^{\frac{dn^d}{2}+d}$ which corresponds to the exact threshold function $f(x)$. \square

H Details of the Proof of Theorem 10

Recall that our proof is by induction on d . For $d = 1$ every matrix $A \in \{0, 1\}^{n \times n}$ is 1-generable. So if we had no extra requirement on the first row, we could just take for the matrix $A^{(1)}$ a matrix constructed by Alon and Vü. To fulfill this last constraint we denote by $C \in \{0, 1\}^{(n-1) \times (n-1)}$ the matrix constructed by Alon and Vü such that $\chi(C) = |\det C_{1,n-1} / \det C| \geq n^{\frac{n}{2}} / 2^{2n-o(n)}$ and let

$$A^{(1)} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & & & \\ 0 & & & \\ 0 & & C & \\ \vdots & & & \\ 0 & & & \end{bmatrix}.$$

It is easy to check that $|\det A_{1,n}^{(1)} / \det A^{(1)}| = |\det C_{1,n-1} / \det C|$.

Now suppose we have a matrix $A^{(d-1)} \in \{0, 1\}^{n^{d-1} \times n^{d-1}}$ satisfying the statement of the theorem. Let us construct a matrix $A^{(d)} \in \{0, 1\}^{n^d \times n^d}$. Note that the number of columns is equal to the number of monomials in M_d , so we need all monomials to label the columns.

Now we define this labeling. We divide the columns in n blocks: the first block consists of the first n^{d-1} columns, the second block of next n^{d-1} columns and so on. Let us denote by $m_1, m_2, \dots, m_{n^{d-1}}$ the monomials corresponding to the columns of $A^{(d-1)}$. We now assign to columns in the j th block the monomials $m_1 x_{d,j}, m_2 x_{d,j}, \dots, m_{n^{d-1}} x_{d,j}$.

Recall that for a matrix $A = \{a_{ij}\}_{i,j=1}^n$ we denote by \bar{A} the matrix obtained by reflecting A in the vertical median, i.e.,

$$\bar{A} = \begin{bmatrix} a_{1,n} & a_{1,n-1} & \dots & a_{1,1} \\ a_{2,n} & a_{2,n-1} & \dots & a_{2,1} \\ \dots & \dots & \dots & \dots \\ a_{n,n} & a_{n,n-1} & \dots & a_{n,1} \end{bmatrix}.$$

Then matrix $A^{(d)}$ is defined by

$$\begin{bmatrix} 1 & 0 & \dots & \boxed{0} & & & & \\ & A^{(d-1)} & & & & & & \\ \hline 0 & \dots & 0 & 1 & \boxed{0} & \dots & 0 & 1 & & & \\ & & & \overline{A}^{(d-1)} & & & & & & & \\ \hline & & & 1 & 0 & \dots & 0 & 1 & 0 & \dots & \boxed{0} & & \\ & & & & & & & A^{(d-1)} & & & & & \\ \hline & & & & & & & & \ddots & & & & \\ & & & & & & & & & & & \ddots & \end{bmatrix}.$$

In the diagonal blocks we have matrices $A^{(d-1)}, \overline{A}^{(d-1)}, A^{(d-1)}, \dots$. On the diagram we have written the first row of each block separately and we have placed a box around an entry for which

$$|\det B / \det A^{(d-1)}| \geq \frac{n^{\frac{1}{2}n^{d-1}}}{2^{2n^{d-1} + o(n^{d-1})}},$$

where by B we denote the corresponding submatrix. In the blocks right below the diagonal only the first row is nonzero and these rows are $e_n^T, e_1^T, e_n^T, \dots$. All other blocks consists exclusively of zeros.

Let us check that matrix $A^{(d)}$ has all required properties (possibly after a permutation of the columns). First of all, the first row is obviously of the correct form. Next we show that it is d -generable. We have already explained how we label the columns. Now we label the rows. Our rows are naturally divided into n blocks and we label each block separately. Let us consider the k th row of the l th block. Let the assignment of values to the variables x_{ij} for $i = 1, \dots, d-1$, $j = 1, \dots, n$ be the same as for the k th row of the matrix $A^{(d-1)}$. If $k > 1$ or $l = 1$ then let $x_{dl} = 1$ and $x_{dj} = 0$ for $j \neq l$. If $k = 1$ and $l > 1$ then let $x_{d,l-1} = x_{d,l} = 1$ and $x_{dj} = 0$ for $j \neq l-1, l$. It is easy to check that with this labeling of columns and rows, each element of the matrix is exactly the value of the monomial corresponding to the column evaluated on the input corresponding to the row.

In the following proof we will make use of Lemma 2.3.1 from [2]. We state it here for the sake of completeness.

Lemma 4 (Alon, Vū). *Let $S = \{s_{i,j}\}_{i,j=1}^{n_1}$ and $T = \{t_{i,j}\}_{i,j=1}^{n_2}$ be two matrices and let*

$$R = \begin{bmatrix} s_{11} & \dots & s_{1n_1} & 0 & \dots & 0 \\ s_{21} & \dots & s_{2n_1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ s_{n_11} & \dots & s_{n_1n_1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 & t_{11} & \dots & t_{1n_2} \\ 0 & 0 & \dots & 0 & t_{21} & \dots & t_{2n_2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & t_{n_21} & \dots & t_{n_2n_2} \end{bmatrix}.$$

Then

$$\chi(R) \geq \left| \frac{\det R_{1, n_1+n_2}}{R} \right| = \left| \frac{\det S_{1n_1} \det T_{1n_2}}{\det S \det T} \right|.$$

□

It only remains to prove the lower bound on $|\det A_{1, n^d}^{(d)} / \det A^{(d)}|$. By repetition of the trick from Lemma 4 we can prove the following claim.

Claim. There exist $k \in \{1, \dots, n^d\}$ such that

$$|\det A_{1, k} / \det A| \geq \frac{n^{\frac{1}{2}n^d}}{2^{2n^d + o(n^d)}}.$$

Proof. We prove by induction on i that for principal submatrix $B^{(i)}$ of $A^{(d)}$ consisting of first i blocks of rows and columns and for some j it is true that

$$\left| \det B_{1, j}^{(i)} / \det B^{(i)} \right| \geq \frac{n^{\frac{1}{2}n^{d-1}i}}{2^{2n^{d-1}i + o(n^{d-1}i)}}.$$

Moreover, we have $j = in^{d-1}$ for odd i and $j = n^{d-1}(i-1) + 1$ for even i . That is, j is the last column of $B^{(i)}$ for odd i and j is the first column of the last block of $B^{(i)}$ for even i .

The basis of induction is obvious. For the inductive step let us consider the matrix $B^{(i)}$ and assume, say, that i is even. Then $B^{(i)}$ has form

$$\left[\begin{array}{ccc|ccc} 1 & 0 & \dots & \boxed{0} & & \\ & & & B^{(i-1)} & & \\ \hline & & 0 & \dots & 0 & 1 \\ & & \boxed{0} & & \dots & 0 & 1 \\ & & & & \overline{A}^{(d-1)} & & \end{array} \right].$$

We know that

$$\left| \det B_{1,(i-1)n^d}^{(i-1)} / \det B^{(i-1)} \right| \geq \frac{n^{\frac{1}{2}n^{d-1}(i-1)}}{2^{2n^{d-1}(i-1)+o(n^{d-1})(i-1)}}$$

and

$$\left| \det \overline{A}_{1,1}^{(d-1)} / \det \overline{A}^{(d-1)} \right| \geq \frac{n^{\frac{1}{2}n^{d-1}}}{2^{2n^{d-1}+o(n^{d-1})}}.$$

From Lemma 4 it immediately follows that one of the entries of the inverse of $B^{(i)}$ is greater than the product of the above two values. More precisely,

$$\left| \det B_{1,(i-1)n^{d+1}}^{(i)} / \det B^{(i)} \right| \geq \frac{n^{\frac{1}{2}n^{d-1}i}}{2^{2n^{d-1}i+o(n^{d-1})i}}.$$

□

Thus if in the matrix $A^{(d)}$ we swap the k th column and the n^d th column we get a matrix as required by Theorem 10. □

I Proof of Proposition 1

This proof is similar to the proof of Lemma 3 from [16]

Recall that we have the equation

$$[B \ B'] \begin{pmatrix} u \\ v \end{pmatrix} = 0.$$

Let us denote by B_1 the matrix consisting of the columns from B' such that in the corresponding monomial there is even number of variables from $(\overline{x}^1 - \overline{y}^1)$. Next denote by B_2 the matrix consisting of columns from B' such that they are not in B_1 and in the corresponding monomial there is even number of variables from $(\overline{x}^2 - \overline{y}^2)$. In general, denote by B_i the matrix consisting of columns from B' such that they are not in B_1, B_2, \dots, B_{i-1} and in the corresponding monomial there is even number of variables from $(\overline{x}^i - \overline{y}^i)$. Note that each column of B' belongs to one of the matrices B_1, \dots, B_n since the monomials corresponding to the columns in B' are not in M_d and have degree at most d . Thus each monomial corresponding to a column in B' does not contain variables from $\overline{x}^i - \overline{y}^i$ for some i . With this notation we can write our system in a more detailed form:

$$[B \ B_1 \ \dots \ B_n] \begin{pmatrix} u \\ v_1 \\ \vdots \\ v_n \end{pmatrix} = 0, \tag{2}$$

where v_1, \dots, v_n are the coordinates of v corresponding to B_1, \dots, B_n respectively.

Now we are going to use a symmetry argument to get rid of the part v in the set of coefficients. Note that each monomial in $M'_d \cup \{\theta\}$ will change sign if we switch \bar{x}^i and \bar{y}^i for some i . The same must thus be true for the polynomial p . This means that if we take values of input corresponding to some row in B and switch values of \bar{x}^i and \bar{y}^i , the value of p will still be 0 and thus the value of f will still be 1. Since the polynomial with coefficients w' represent the function f the value of this polynomial should also be 0 on this input. Applying this to $i = 1$ and all rows of the matrix we get

$$[-B \ B_1 \ -B_2 \ \dots \ -B_n] \begin{pmatrix} u \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = 0$$

since all monomials corresponding to columns which are not in B_1 contains odd number of variables from $(\bar{x}^1 - \bar{y}^1)$. Subtracting this system from (2) and dividing by 2 we get

$$[B \ 0 \ B_2 \ \dots \ B_n] \begin{pmatrix} u \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = 0$$

which is equal to

$$[B \ B_2 \ \dots \ B_n] \begin{pmatrix} u \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = 0$$

Arguing by the same way consecutively for $(\bar{x}^2 - \bar{y}^2), \dots, (\bar{x}^d - \bar{y}^d)$ we get the system

$$Bu = 0.$$

□

J Lower bound on $\max w_d^T$

In this section we will prove the following theorem.

Theorem 12.

$$\max w_d^T(2dn) \geq \frac{n^{\frac{1}{2}n^d}}{2^{2n^d + o(n^d)} + 2d}.$$

Proof. As in the proof of Theorem 11 we have variables $x_{ij}, y_{ij} \in \{0, 1\}$ for $i = 1, \dots, d$ and $j = 1, \dots, n$ and we consider polynomials in variables $x - y$ and

$x + y$. As in the proof of Theorem 11 we consider matrix $A^{(d)}$ and, again, in the labeling of $A^{(d)}$ we substitute each variable x_{ij} by $x_{ij} - y_{ij}$.

Let

$$\alpha_i = \text{sign}(-1)^i \det A_{i,n^d}^{(d)} \det A^{(d)}$$

for $i = 1, \dots, n^d$ and let vector w be a solution of the system

$$A^{(d)}w = \alpha.$$

Let $f(x, y)$ be a polynomial threshold function of degree d represented by a polynomial w .

Now consider any integer representation w' of this function. We have

$$[A^{(d)} \ B'] \begin{pmatrix} u \\ v \end{pmatrix} = \beta, \quad (3)$$

where B', u, v mean the same as in the proof of Theorem 11 and β is an integer vector (since w' is an integer vector) such that $\text{sign } \beta_i = \alpha_i$.

By symmetry argument analogous to one in the proof of Theorem 11 we can get

$$A^{(d)}u = \beta',$$

where β'_i are rational numbers with denominators 2^d and $\text{sign } \beta'_i = \alpha_i$. Now if we consider this equation as a system of linear equations with variables u , by Cramer's rule we get

$$|w_{n^d}| = \left| \sum_{i=1}^{n^d} (-1)^i \frac{\beta'_i A_{i,n^d}^{(d)}}{A^{(d)}} \right| = \sum_{i=1}^{n^d} \left| \frac{\beta'_i A_{i,n^d}^{(d)}}{A^{(d)}} \right| \geq \left| \frac{\beta'_1 A_{1,n^d}^{(d)}}{A^{(d)}} \right| \geq \frac{n^{\frac{1}{2}n^d}}{2^{2n^d + o(n^d) + 2d}},$$

where we have guaranteed the second equation by the choice of α . \square

Appendix References

20. Aliev, I.: Siegel's lemma and sum-distinct sets. *Discrete and Computational Geometry* 39(1-3), 59–66 (2008)
21. Bohman, T.: A construction for sets of integers with distinct subset sums. *Electronic Journal of Combinatorics* 5, 1–14 (1998)
22. Conway, J., Guy, R.: Sets of natural numbers with distinct subset sums. *Notices of the American Mathematical Society* 15, 345 (1968)
23. Elkies, N.D.: An improved lower bound on the greatest element of a sum-distinct set of fixed order. *Journal of Combinatorial Theory, Series A* 41(1), 89–94 (1986)
24. Erdős, P.: Problems and results in additive number theory. In: *Colloque sur la Théorie des Nombres*, Bruxelles, 1955, pp. 127–137. George Thone, Liège (1956)
25. Williamson, J.: Determinants whose elements are 0 and 1. *American Mathematical Monthly* 53(8), 427–434 (1946)