Anti-Hadamard matrices, coin weighing, threshold gates and indecomposable hypergraphs

Noga Alon *   Văn H. Vũ †

February 22, 2002

Abstract

Let $\chi_1(n)$ denote the maximum possible absolute value of an entry of the inverse of an n by n invertible matrix with $0,1$ entries. It is proved that $\chi_1(n) = n^{(\frac{1}{2}+o(1))n}$. This solves a problem of Graham and Sloane.

Let $m(n)$ denote the maximum possible number m such that given a set of m coins out of a collection of coins of two unknown distinct weights, one can decide if all the coins have the same weight or not using n weighings in a regular balance beam. It is shown that $m(n) = n^{(\frac{1}{2}+o(1))n}$. This settles a problem of Kozlov and Vũ.

Let $D(n)$ denote the maximum possible degree of a regular multi-hypergraph on n vertices that contains no proper regular nonempty subhypergraph. It is shown that $D(n) = n^{(\frac{1}{2}+o(1))n}$. This improves estimates of Shapley, van Lint and Pollak.

All these results and several related ones are proved by a similar technique whose main ingredient is an extension of a construction of Håstad of threshold gates that require large weights.

1 Introduction

For a real matrix A, the *spectral norm* of A is defined by $\|A\|_s = \sup_{x\neq 0} |Ax|/|x|$. If A is invertible, the *condition number* of A is $c(A) = \|A\|_s\|A^{-1}\|_s$. This quantity measures the sensibility of the equation $Ax = b$ when the right hand side is changed. If $c(A)$ is large, then A is called *ill-conditioned.* For the above reason, ill-conditioned matrices are important in numerical algebra, and have been studied extensively by various researchers (see, e.g., [7], [16] and their references). In [10], Graham and Sloane consider the special case of ill-conditioned matrices, whose entries lie in the set $\{0, 1\}$ or in the set $\{-1, 1\}$. These special cases are of interest not only in linear algebra, since $(0, 1)$ and $(-1, 1)$ matrices are basic objects in combinatorics and related areas. In their paper Graham and Sloane study the most ill-conditioned $(0, 1)$ (or $(-1, 1)$) matrices, which they call *anti-Hadamard* matrices.

*AT & T Research, Murray Hill, NJ 07974, USA and Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email address: noga@math.tau.ac.il.
Research supported in part by a USA Israeli BSF grant and by the Fund for Basic Research administered by the Israel Academy of Sciences.

†Department of Mathematics, Yale University, 10 Hillhouse Ave., New Haven, CT-06520, USA. Email address: vuha@math.yale.edu

For matrices with such restricted entries, many quantities are equivalent to the condition number. Let A be a non-singular $(0,1)$ matrix and put $B = A^{-1} = (b_{ij})$. The following quantities are considered in [10], where in both cases the maximum is taken over all invertible n by n matrices with $0,1$ entries.

- $\chi(A) = \max_{i,j} |b_{ij}|$ and $\chi(n) = \max_A \chi(A)$
- $\mu(A) = \sum_{i,j} b_{ij}^2$ and $\mu(n) = \max_A \mu(A)$.

It is shown in [10] that $c(2.274)^n \le \chi(n) \le 2(n/4)^{n/2}$ for some absolute positive constant c, and consequently that $c^2(5.172)^n \le \mu(n) \le 4n^2(n/4)^n$, and the authors raise the natural problem of closing the gap between these bounds.

Our first result here determines the asymptotic behaviour of $\chi(n)$, as well as that of the analogous quantity for $(-1,1)$-matrices. It turns out that this function is $n^{(\frac{1}{2}+o(1))n}$ in both cases, where the $o(1)$ term tends to 0 as n tends to infinity. This implies that the maximum possible condition numbers of such n by n matrices is also $n^{(\frac{1}{2}+o(1))n}$.

Our lower-bound is by an explicit construction of appropriate ill conditioned matrices. This construction is based on a (modified version of) a construction of Håstad [11] and an extension of it.

It turns out that this result has many interesting applications to several seemingly unrelated problems, listed below.

- *Flat simplices:* We show that the minimum possible positive distance between a vertex and the opposite facet in a nontrivial simplex determined by $(0,1)$ vectors in R^n is $n^{-(\frac{1}{2}+o(1))n}$. This answers another question suggested in [10].

- *Threshold gates with large weights:* A *threshold gate* of n inputs is a function $F : \{-1,1\}^n \mapsto \{-1,1\}$ defined by

$$F(x_1,\ldots,x_n) = sign(\sum_{i=1}^{n} w_i x_i - t),$$

  where $w_1,\ldots,w_n,t$ are reals called *weights*, chosen in such a way that the sum $\sum_{i=1}^n w_i x_i - t$ is never zero for $(x_1,\ldots,x_n) \in \{-1,1\}^n$. Threshold gates are the basic building blocks of Neural Networks, and have been studied extensively. See, e.g., [12] and its references. It is easy to see that every threshold gate can be realized with integer weights. Various researchers proved that there is always a realization with integer weights satisfying $|w_i| \le n^{(\frac{1}{2}+o(1))n}$, and Håstad [11] proved that this is tight (up to the $o(1)$ term) for all values of n which are powers of 2. Here we extend his construction and show that this upper bound is tight for all values of n.

- *Coin weighing:* Let $m(n)$ denote the maximum possible number m such that given a set of m coins out of a collection of coins of two unknown distinct weights, one can decide if all the coins have the same weight or not using n weighings in a regular balance beam. We prove that $m(n) = n^{(\frac{1}{2}+o(1))n}$. This is tight up to the $o(1)$-term and settles a problem of Kozlov and Vũ [14]. A similar estimate holds when there are more potential weights , but they satisfy a certain generic assumption, and even when there is no assumption on the possible weights of the coins, but there is a given coin which is known to be either the heaviest or the lightest among the given coins.

- *Indecomposable hypergraphs:* An (multi-) hypergraph is *indecomposable* if it is regular, but none of its proper subhypergraphs is regular. Let $D(n)$ be the maximum possible degree of an indecomposable hypergraph on n points. The problem of estimating $D(n)$ is motivated by questions in Game Theory and has been considered by many researchers (see [8] for a survey). Here we show that $D(n) = n^{(\frac{1}{2}+o(1))n}$.

All problems above are closely related, and the lower-bounds for all of them are obtained by applying an appropriate ill-conditioned $(0,1)$ or $(-1,1)$ matrix. All the upper-bounds rely on Hadamard inequality, which is the following well known fact.

Lemma 1.1. *If A is a matrix of order n, then $|\det A| \le \prod_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}^2)^{1/2}$, where a_{ij} is the entry in row i and column j.* $\Box$

The rest of this paper is organized as follows. In the rest of this section we introduce some (mostly standard) notation. In Section 2 we construct ill conditioned matrices with $(0,1)$ entries and with $(-1,1)$ entries. Section 3 contains the proofs of all the above mentioned applications and the final section 4 contains some concluding remarks and open problems.

Notation.

For a matrix B, b_{ij} denotes the entry in row i and column j, and B_{ij} denotes the submatrix obtained from B by deleting the row i and column j. J_n and I_n are the all-one and the identity matrix of order n, respectively. Ill-conditioned matrices are always non-singular square matrices. The *direct sum* of two square matrices A and B is $A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$. The coordinates of a vector x of length n are denoted by lower-indexed letters $x_1, x_2, \ldots, x_n$, and x is written in the form $x = (x_1, x_2, \ldots, x_n)$, or sometimes in the form $x = (x_i)_{i=1}^n$. We denote by $\mathbf{1}_n$ the all-one vector of length n. The l_1 and l_∞ norms of x are $\|x\|_1 = \sum_{i=1}^n |x_i|$ and $\|x\|_\infty = \max_{i=1}^n |x_i|$, respectively. A vector is *integral* if all of its coordinates are integers. $\{0,1\}^n$ and $\{-1,1\}^n$ denote the sets of all vectors of length n, with coordinates from the sets $\{0,1\}$ and $\{-1,1\}$, respectively. It is convenient to note that each of these sets is the set of vertices of the corresponding hypercube in $\mathbf{R}^n$.

As usual, $\theta(n)$ represents a quantity satisfying $c_1 n \le \theta(n) \le c_2 n$, where $0 < c_1 < c_2$ are constants. Since most results in terms of n in this paper are asymptotic, we always assume that n is sufficiently large, whenever this is needed. All logarithms used in the paper are in base 2. A real function f is called *super-multiplicative* if it satisfies $f(m+n) \ge f(m)f(n)$ for all admissible m, n.

In the proofs we apply the following simple and well-known elementary equalities, whose proofs are omitted.

Lemma 1.2.

(1) For any positive integer m: $\sum_{k=0}^{m} \binom{m}{k} = 2^m$ and $\sum_{k=1}^{m} k\binom{m}{k} = m2^{m-1}$

(2) $\sum_{i=1}^{\infty} i(2^{1-i}) = 4$. $\Box$

2 Ill conditioned matrices

The purpose of this section is to estimate the maximum possible condition numbers of $(0,1)$ and $(-1,1)$ matrices. First, let us introduce some notation. Let $\mathcal{A}_n^1$ and $\mathcal{A}_n^2$ denote the sets of invertible $(0,1)$ and $(-1,1)$ matrices of order n, respectively. For an invertible matrix A, let $\chi(A)$ denote the maximum absolute value of an entry of A^{-1}. It is easy to see that $\chi(A)$ is invariant under permutations and sign changes of rows and columns of A. Though this is not true for arbitrary matrices, it will be shown in subsection 3.1 that the condition numbers of $(0,1)$ and $(-1,1)$ matrices A have the same order of magnitude as $\chi(A)$; large χ implies that the condition number is large, and thus that the matrix is very ill-conditioned. Thus we use here $\chi(A)$ to measure how ill-conditioned the matrix A is.

Define $\chi_i(n) = \max_{A \in \mathcal{A}_n^i} \chi(A)$, where $i = 1, 2$. The following theorem determines the asymptotic behaviour of $\chi_i(n)$. Since all the results in Section 3 are based on this theorem, we call it the main theorem. We emphasize in the second part of the theorem that the lower-bound is constructive; this will play a role in the applications.

The Main Theorem. *For $i = 1, 2$,*

1. *The functions $\chi_i(n)$ are super-multiplicative and satisfy*

$$2^{\frac{1}{2}n\log n - n(1+o(1))} \geq \chi_i(n) \geq 2^{\frac{1}{2}n\log n - n(2+o(1))}.$$

2. *One can construct explicitly a matrix $C^i \in \mathcal{A}_n^i$ such that*

$$\chi(C^i) \geq 2^{\frac{1}{2}n\log n - n(2+o(1))}.$$

By *explicit construction* we mean here the existence of an algorithm that constructs, given n, an n by n matrix satisfying the above inequality in time which is polynomial in n.

The upper-bound for $\chi_2(n)$ is quite easy. Consider $A \in \mathcal{A}_n^2$, and let b_{ij} be an element of $B = A^{-1}$. By Cramer's rule $b_{ij} = (-1)^{i+j} \det A_{ij} / \det A$, thus $|b_{ij}| = |\det A_{ij} / \det A|$.

Since A_{ij} is a $(-1,1)$ matrix of order $(n-1)$, by Hadamard inequality $\det A_{ij} \leq (n-1)^{(n-1)/2} = 2^{\frac{1}{2}n\log n - o(n)}$. On the other hand, $|\det A|$ is at least 2^{n-1}. To see this, one can add the first row of A to each other row, thus getting rows with $2, 0$ and -2 entries. Thus, the determinant of A is divisible by 2^{n-1}, and hence $|\det A| \geq 2^{n-1}$. This implies that $|b_{ij}| \leq 2^{\frac{1}{2}n\log n - n(1+o(1))}$.

The proof of the Main Theorem will be presented in the following steps. In subsection 2.1 we construct a matrix $A \in \mathcal{A}_n^2$ for $n = 2^m$, such that $\chi(A)$ differs from the upper-bound by a sub-exponential factor only. This construction is based on the ideas of Håstad in [11]. However, our construction is somewhat simpler and the proof of its properties is slightly more direct than that given in [11].

In subsection 2.2 we describe a simple, known connection between the two classes $\mathcal{A}_{n-1}^1$ and $\mathcal{A}_n^2$. Using this, we obtain the upper-bound for $\chi_1(n)$, as well as $(0,1)$ matrices of orders $n = 2^m - 1$ with large χ. In subsection 2.3 we establish the super-multiplicativity of $\chi_i(n)$. We complete the proof of the theorem in subsection 2.4, where we construct $(0,1)$ and $(-1,1)$ matrices of arbitrary order n, for which the lower-bound holds, by combining the supermultiplicativity with the constructions for powers of 2.

2.1 Ill-conditioned $(-1,1)$ matrices of order 2^m

Theorem 2.1.1 *For $n = 2^m$ there is a matrix $A \in \mathcal{A}_n^2$ such that*

$$\chi(A) = 2^{\frac{1}{2}n\log n - n(1+o(1))}.$$

Proof. The matrix A is constructed explicitly as follows. Let Ω be a set of m elements. Order the subsets α_i, $i = 1, .., 2^m$ of Ω in such the way that $|\alpha_i| \leq |\alpha_{i+1}|$ and $|\alpha_i \triangle \alpha_{i+1}| \leq 2$, where $|\alpha|$ denotes the cardinality of α and $\alpha \triangle \beta$ denotes the symmetric difference between the two sets α and β. To achieve such an ordering, it suffices to order all the subsets of the same cardinality, and this can be easily done by induction. For a detailed proof, we refer to Lemma 2.1 in [11]. It is convenient to let α_0 denote the empty set. Our matrix A is defined by the following simple rules. For every $1 \leq i, j \leq n$:

(1) If $\alpha_j \cap (\alpha_{i-1} \cup \alpha_i) = \alpha_{i-1} \triangle \alpha_i$ and $|\alpha_{i-1} \triangle \alpha_i| = 2$, then $a_{ij} = -1$.

(2) If $\alpha_j \cap (\alpha_{i-1} \cup \alpha_i) \neq \emptyset$ but (1) does not occur, then $a_{ij} = (-1)^{|\alpha_{i-1} \cap \alpha_j| + 1}$.

(3) If $\alpha_j \cap (\alpha_{i-1} \cup \alpha_i) = \emptyset$, then $a_{ij} = 1$.

We next prove that A has the required property.

Let Q be the n by n matrix given by $q_{ij} = (-1)^{|\alpha_i \cap \alpha_j|}$. It is easy and well known that Q is a symmetric Hadamard matrix, that is $Q^2 = nI_n$. Next, we construct a matrix L row by row as follows. For the i^{th} row of L $(i > 1)$, consider the set α_i. define $A_i = \alpha_{i-1} \cup \alpha_i$. Define also $F_i = \{\alpha_s | \alpha_s \subset A_i, |\alpha_s \cap (\alpha_{i-1} \triangle \alpha_i)| = 1\}$ if $|\alpha_{i-1} \triangle \alpha_i| = 2$ and $F_i = \{\alpha_s | \alpha_s \subset A_i\}$ if $|\alpha_{i-1} \triangle \alpha_i| = 1$. Note that if $|\alpha_i| = k$, then $|F_i| = 2^k$ in both cases.

Set $l_{ij} = 0$ iff $\alpha_j \notin F_i$. Among the remaining 2^k entries of the row, let $l_{i,i-1} = \frac{1}{2}^{k-1} - 1$, and let all others be $1/2^{k-1}$. By the property of the ordering, it is clear that if $j > i$ then $\alpha_j \notin F_i$. For $i = 1$, $a_{11} = 1$ is the only non-zero element of the first row. Thus L is a lower triangular matrix.

Lemma 2.1.2 *A has the following factorization: $A = LQ$.*

Proof. Consider the inner product of the i^{th} row of L and the j^{th} column of Q

$$\sum_{s=1}^{n} l_{is} q_{sj} = \sum_{s, l_{is} \neq 0} (1/2^{k-1})(-1)^{|\alpha_s \cap \alpha_j|} + (-1)(-1)^{|\alpha_{i-1} \cap \alpha_j|}$$

$$= (1/2^{k-1}) \sum_{\alpha_s \in F_i} (-1)^{|\alpha_s \cap \alpha_j|} + (-1)^{|\alpha_{i-1} \cap \alpha_j| + 1} = \Sigma_{ij} + (-1)^{|\alpha_{i-1} \cap \alpha_j| + 1}.$$

Consider three subcases according to the definition of A. If (1) occurs, then each term in Σ_{ij} is $-1/2^{k-1}$, so $\Sigma_{ij} = -2$. Moreover, the second summand is 1 so the inner product is -1. If (2) occurs, then by symmetry, half of the members of F_i have an odd (even) intersection with α_j, so half of the terms in Σ_{ij} are $-1/2^{k-1}$, and hence $\Sigma_{ij} = 0$ and the inner product is equal to the second summand. Finally, if (3) occurs, all the terms in Σ_{ij} are $1/2^{k-1}$ and $\Sigma_{ij} = 2$, the second summand is -1, and thus the product is 1. This proves the Lemma.$\square$

Let i_0 be the first index such that α_{i_0} has three elements. Let δ be the $(0,1)$ vector of length n, in which i_0 is the only non-zero coordinate. Consider the equation $Lx = \delta$. For $i > 1$, its i^{th} row equation reads

$$\sum_{\alpha_j \in F_i} (1/2^{k-1})x_j - x_{i-1} = \delta_i$$

or equivalently,

$$x_i = (2^{k-1} - 1)x_{i-1} - \sum_{\alpha_j \in F_i \setminus \{\alpha_{i-1}, \alpha_i\}} x_j + 2^{k-1}\delta_i$$

Observe that for $i < i_0$, $\delta_i = 0$, thus $x_i = 0$. Furthermore, $x_{i_0} = 2^{3-1}\delta_{i_0} = 4$ and $x_{i_0+1} = (2^2 - 1)x_{i_0} = 3x_{i_0}$. By induction we next show that $|x_i| > (2^{k-1} - 2)|x_{i-1}|$ for $i > i_0$. Indeed if the statement holds for $i - 1$ then $|x_{i-1}| > 2|x_{i-2}| > 4|x_{i-3}|...$, hence the above sum of the elements x_j is majorized by the sum $\sum_{t=1}^{\infty}(1/2^t)|x_{i-1}| = |x_{i-1}|$. Thus we have,

$$|x_i| \geq (2^{k-1} - 2)|x_{i-1}| + |x_{i-1}| - \sum_{\alpha_j \in F_i \setminus \{\alpha_{i-1}, \alpha_i\}} |x_j| > (2^{k-1} - 2)|x_{i-1}|.$$

This proves the statement for i, completing the induction.

One can deduce from here that all the numbers x_i are non-negative. By the statement just proved it follows that:

$$x_n > \prod_{k=3}^{m} (2^{k-1} - 2)^{\binom{m}{k}} = \prod_{k=3}^{m} 2^{(k-1)\binom{m}{k}} \prod_{k=3}^{m} (1 - \frac{2}{2^{k-1}})^{\binom{m}{k}}$$

Using the equalities in Lemma 1.2, the first product is

$$2^{\sum_{k=1}^{m}(k-1)\binom{m}{k} - \binom{m}{2}} = 2^{m2^{m-1} - 2^m + 1 - O(m^2)}$$

$$= 2^{(1/2)n\log n - n - O(\log^2 n)} = 2^{(1/2)n\log n - n(1+o(1))}$$

The reader can verify that the second product is at least $2^{-o(n)}$. In fact it can be lower-bounded by $e^{-n^{\beta}} = 2^{-o(n)}$, for some $\beta < 1$. This can be done by observing that $1 - x > e^{-2x}$ for $x < 1/2$ and by some simple manipulations (see [11] for the detailed computation.) Thus we have $x_n \geq 2^{(1/2)n\log n - n(1+o(1))}$.

We complete the proof by considering the equation $Ay = \delta$. By Cramer's rule $|y_i| = |\det A_{i_0 j}/\det A|$. On the other hand, $A = LQ$, so $Qy = x$ or $y = Q^{-1}x$. As mentioned in the beginning of the proof $Q^{-1} = (1/n)Q$, thus we have $y = (1/n)Qx$ or equivalently $y_i = 1/n\sum_j q_{ij}x_j$. Since $|q_{ij}| = 1$, and since $x_n > 4x_{n-1} > 8x_{n-2} > ...$ we conclude that $|y_i| > (1/n)(1/2)x_n$. Therefore $|y_i| = 2^{\frac{1}{2}n\log n - n(1+o(1))}$. In other words, all the elements of the i_0^{th} column of A^{-1} have the required order of magnitude.

If one chooses any $j_0 > i_0$ so that the product $\prod_{k=|\alpha_{j_0}|}^{m}(2^{k-1} - 2)^{\binom{m}{k}}$ has order of magnitude $2^{\frac{1}{2}n\log n - \theta(n)}$, then the corresponding terms $\det A_{i_0 j}/\det A$ also have this order of magnitude. This shows that A^{-1} has, in fact, many columns consisting of large entries. $\square$

Remark. The matrix A constructed above has minimal determinant $\det A = 2^{n-1}$. Indeed, observe that $\det A = \det L \det Q$. Moreover, $\det Q = n^{n/2} = 2^{m2^{m-1}}$, since Q is a Hadamard matrix. Furthermore, L is lower-triangular, implying that

$$\det L = \prod_{i=1}^{n} l_{ii} = \prod_{k=1}^{m} 2^{-(k-1)\binom{m}{k}} = 2^{(2^m-1)-m2^{m-1}}$$

This yields $\det A = \det L \det Q = 2^{2^m-1} = 2^{n-1}$.

2.2 The connection between $\mathcal{A}^1_{n-1}$ and $\mathcal{A}^2_n$

In this subsection we describe a simple connection between the two classes $\mathcal{A}^1_{n-1}$ and $\mathcal{A}^2_n$. Consider the map Φ which assigns to any matrix $B \in \mathcal{A}^1_{n-1}$ a matrix $\Phi(B) \in \mathcal{A}^2_n$ in the following way:

$$\Phi(B) = \begin{pmatrix} 1 & \mathbf{1}_{n-1} \\ -\mathbf{1^T}_{n-1} & 2B - J_{n-1} \end{pmatrix}$$

This map has a nice and simple geometric interpretation. Let P_i be the point in $\mathbf{R}^{n-1}$ represented by the i^{th} row of B, $i = 1, 2, \ldots, n-1$. Similarly, let Q_i be the point in $\mathbf{R}^n$ represented by the $(i+1)^{th}$ row of $\Phi(B)$, for $i = 0, 1, \ldots, n-1$. Now identify the unit hypercube of $\mathbf{R}^{n-1}$ with the unit hypercube of the hyperplane $x_1 = 0$ in $\mathbf{R}^n$. Then P_i will be identified with the midpoint of the segment Q_0Q_i.

The above map is clearly invertible, and by simple row operations (see [6]) it follows that $|\det \Phi(B)| = 2^{n-1}|\det B|$. If B is invertible, so is $\Phi(B)$, and

$$\Phi(B)^{-1} = \begin{pmatrix} 1 - \frac{1}{2}\mathbf{1}_{n-1}B^{-1}\mathbf{1^T}_{n-1} & -\frac{1}{2}\mathbf{1}_{n-1}B^{-1} \\ \frac{1}{2}B^{-1}\mathbf{1^T}_{n-1} & \frac{1}{2}B^{-1} \end{pmatrix}$$

Moreover, note that every matrix in $\mathcal{A}^2_n$ can be normalized to have the first column and row like those in a typical $\Phi(B)$; all one has to do is to multiply some rows and columns by -1, if needed. Thus, in a loose sense, Φ is a bijection. Multiply all the rows of the matrix A constructed in subsection 2.1, except the first one, by -1 to get a matrix A_1 whose first column is $(1, -1, -1, .., -1)$ and whose first row is the all 1 vector. Therefore, there is a $(0, 1)$ matrix A' of order $(n-1)$ such that $\Phi(A') = A_1$.

By the above formula for $\Phi(B)^{-1}$, for every entry of A_1^{-1} which is not in the first row or in the first column, the corresponding entry of A'^{-1} has the same absolute value up to a factor of 2.

By the discussion in subsection 2.1, we know that A_1^{-1} contains many columns of large entries (and in particular the i_0^{th} column). It follows that A'^{-1} also has many columns of large entries, and $\chi(A') = 2^{\frac{1}{2}n\log n - n(1+o(1))}$. The formula of $\Phi(B)^{-1}$ also proves the upper-bound for $\chi_1(n)$, as a consequence of the upper-bound for $\chi_2(n)$.

Corollary 2.2.1. *For every $n = 2^m - 1$ there is a matrix $A' \in \mathcal{A}^1_n$ such that $\chi(A') \geq 2^{\frac{1}{2}n\log n - n(1+o(1))}$.*

The matrix $\mathbf{1}_1 \oplus A'$ is of order $n + 1 = 2^m$ and satisfies $\chi(\mathbf{1}_1 \oplus A') = \chi(A')$. Since it will be more convenient to use matrices of order power of 2 in subsection 2.4, we reformulate the last corollary as follows

Corollary 2.2.2. *For every n which is a power of 2 there is a matrix $A' \in \mathcal{A}^1_n$ such that*

$$\chi(A') \geq 2^{\frac{1}{2}n\log n - n(1+o(1))}.$$

Note that since we are interested in asymptotic formulas, there is no difference between n and $n+1$

Remark. Since A and A_1 have determinants with minimum possible absolute value, $\det A = -\det A_1 = 2^{n-1}$, A' also has a determinant with minimum possible absolute value, $|\det A'| = 1$, by the property of the map Φ.

2.3 The super-multiplicativity of $\chi_i(n)$

We first prove that $\chi_1(n)$ is super-multiplicative. To this end, it suffices to show that for any two matrices $S \in \mathcal{A}^1_{n_1}$ and $T \in \mathcal{A}^1_{n_2}$, there is a matrix $R \in \mathcal{A}^1_{n_1+n_2}$, such that $\chi_1(R) \geq \chi_1(S)\chi_1(T)$. The main ingredient in the proof of this fact is the following operation, denoted by $\diamond$, which glues S and T together.

Let S and T be two non-singular matrices of orders n_1 and n_2, respectively. We define $S \diamond T$ as follows. First rearrange the rows and columns of S and T in such a way that $\chi(S) = |\det S_{1n_1}/\det S|$ and $\chi(T) = |\det T_{1n_2}/\det T|$. Suppose now that S and T have this property, then $R = S \diamond T$ has order $n_1 + n_2$ and is obtained from $S \oplus T$ by switching the element r_{n_1+1,n_1} from zero to one. Therefore, R looks as follows:

$$R = \begin{bmatrix} s_{11} & \dots & s_{1n_1} & 0 & \dots & 0 \\ s_{21} & \dots & s_{2n_1} & 0 & \dots & 0 \\ . & \dots & . & . & \dots & . \\ . & \dots & . & . & \dots & . \\ s_{n_11} & \dots & s_{n_1n_1} & 0 & \dots & 0 \\ 0 & 0\dots0 & 1 & t_{11} & \dots & t_{1n_2} \\ 0 & 0\dots0 & 0 & t_{21} & \dots & t_{2n_2} \\ . & \dots & . & . & \dots & . \\ . & \dots & . & . & \dots & . \\ 0 & 0\dots0 & 0 & t_{n_21} & \dots & t_{n_2n_2} \end{bmatrix}$$

The following Lemma shows that R has the required property.

Lemma 2.3.1 $\chi(S \diamond T) \geq \chi(S)\chi(T)$

Proof First we need the following notion. A matrix M is called *near lower-triangular* if it has the form $\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$, where A and B are square matrices. Similarly, M is *near upper-triangular* if it has the form $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$

Obviously, if M is either near lower-triangular or near upper-triangular as above, then $\det M = \det A \det B$.

Consider the matrix $R = S \diamond T$. It has order $n = n_1 + n_2$. By the construction, R is a near lower-triangular matrix of the form $\begin{pmatrix} S & 0 \\ C & T \end{pmatrix}$. Thus, $\det R = \det S \det T$. Furthermore, consider the submatrix R_{1,n_1+n_2} of R. Again by the construction, this has a near upper-triangular form $\begin{pmatrix} S' & D \\ 0 & T' \end{pmatrix}$, where S' is the submatrix S_{1n_1} of S, and T' is

obtained from T by deleting its last column and by adding a column $(1, 0, \ldots, 0)$ to its left. Since the first column of T' has only one non-zero element $t'_{11} = 1$, it is clear that $\det T' = \det T'_{11} = \det T_{1n_2}$. Hence $\det R_{1,n_1+n_2} = \det S' \det T' = \det S_{1n_1} \det T_{1n_2}$.
To conclude the proof of the Lemma observe that

$$\chi(R) \geq |\frac{\det R_{1n}}{\det R}| = |\frac{\det S_{11} \det T_{11}}{\det S \det T}| = \chi(S)\chi(T),$$

as needed.
We can use a similar idea to prove the super-multiplicativity of $\chi_2(n)$. In fact, χ_2 satisfies a stronger inequality: $\chi_2(n_1 + n_2 - 1) \geq 2\chi_2(n_1)\chi_2(n_2)$. The glueing operation in this case is a little more technical. Consider two $(-1, 1)$ matrices S and T of sizes n_1 and n_2, respectively. By changing signs of columns and rows, we can suppose that every element of the last column and the last row of S is $(1, 1, \ldots, 1)$, the first row of T is $(1, 1, \ldots, 1)$ and the first column of T is $(1, -1, 1, \ldots, 1)$ (the second coordinate of the last vector is the only -1). Moreover, we can suppose that $\chi(S) = |\det S_{1n_1}/\det S|$ and $\chi(T) = |\det T_{2n_2}/\det T|$.

Now consider the matrix R of order $n = n_1 + n_2 - 1$ which has S as its $(1, 2, \ldots, n_1)$ principal submatrix, and T as its $(n_1, n_1 + 1, \ldots, n_1 + n_2 - 1)$ principal submatrix, and all non-defined entries are 1. By subtracting the n_1^{th} row from the rows $1, 2, \ldots, n_1 - 1$ one can prove that $|\det R| = |\det S \det T|$. Furthermore, by subtracting the same row from rows $n_1 + 1, \ldots, n_1 + n_2 - 1$ one can show that $|\det R_{1n}| = 2|\det S_{1n_1} \det T_{2n_2}|$. This proves the desired inequality. The (simple) details are left to the reader. □

2.4 Ill-conditioned matrices of arbitrary order

Let n be a large positive integer. We construct a matrix C in $\mathcal{A}_n^1$ which satisfies $\chi(C) \geq 2^{\frac{1}{2}n\log n - n(2+o(1))}$.
Write n as a sum of powers of 2, $n = \sum_{i=1}^r 2^{q_i}$, where $q_1 > q_2 > \ldots > q_r \geq 0$. Let $n_i = 2^{q_i}$. Let A_i be an ill-conditioned matrix of order n_i constructed in subsection 2.2 which satisfies $\chi(A_i) = 2^{\frac{1}{2}n_i \log n_i - n_i(1+o(1))}$. Consider the $(0, 1)$ matrix $C = A_1 \diamond (A_2 \diamond (\ldots (A_{r-1} \diamond A_r)) \ldots)$. By the definition of the operation $\diamond$, C has order $\sum_{i=1}^r n_i = n$. To estimate $\chi(C)$ we apply Lemma 2.3.1 and conclude that

$$\chi(C) \geq \prod_{i=1}^r \chi(A_i) = 2^{\sum_{i=1}^r \frac{1}{2}n_i \log n_i - \sum_{i=1}^r n_i(1+o(1))}$$

In order to estimate the right hand side properly, we need the following Lemma:

Lemma 2.4.2. *If $q_1 > q_2 > \ldots > q_r \geq 0$ are integers, and $n_i = 2^{q_i}$, $N = \sum_{i=1}^r n_i$ then*

$$\zeta(N) = \frac{1}{N}(\sum_{i=1}^r n_i \log N - \sum_{i=1}^r n_i \log n_i) \leq 2$$

Proof. We call the set $\Upsilon = \{q_1, q_2, \ldots, q_r\}$ *full* if it contains all non-negative integers not larger than q_1. The proof follows from the following two facts.
Fact 1. If Υ is full, then $\zeta(N) \leq 2$.
Fact 2. If Υ is not full, q is a non-negative integer less than q_1 not in Υ, and $n_* = 2^q$, then $\zeta(N + n_*) \geq \zeta(N)$.

Fact 1 is straightforward. We prove Fact 2. First, we rewrite ζ in a more convenient form,

$$\begin{aligned}\zeta(N) &= \sum_{i=1}^{r} \frac{n_i}{N} \log \frac{N}{n_i} \\ &= \sum_{i=1}^{r} \frac{n_i}{N} \log \frac{N}{n_1}\frac{n_1}{n_i} \\ &= \sum_{i=1}^{r} \frac{n_i}{N} \log \frac{n_1}{n_i} + \log \frac{N}{n_1}\end{aligned}$$

By this, we have,

$$\zeta(N+n_*) = \sum_{i=1}^{r} \frac{n_i}{N+n_*} \log \frac{n_1}{n_i} + \frac{n_*}{N+n_*} \log \frac{n_1}{n_*} + \log \frac{N+n_*}{n_1}$$

Hence,

$$\zeta(N+n_*) - \zeta(N) = \frac{n_*}{N+n_*} \log \frac{n_1}{n_*} + \log \frac{N+n_*}{N} - \sum_{i=1}^{r} \frac{n_* n_i}{N(N+n_*)} \log \frac{n_1}{n_i}$$

We prove $\zeta(N+n_*) > \zeta(N)$ by showing that in fact,

$$\frac{n_*}{N+n_*} \log \frac{n_1}{n_*} - \sum_{i=1}^{r} \frac{n_* n_i}{N(N+n_*)} \log \frac{n_1}{n_i} > 0$$

By a simplification and a rearrangement, this is equivalent to

$$N \log \frac{n_1}{n_*} > \sum_{i=1}^{r} n_i \log \frac{n_1}{n_i}$$

Since $N = \sum_{i=1}^{r} n_i$, the last inequality is equivalent, after some simplification and rearrangement of terms, to

$$\sum_{i=1}^{r} n_i \log \frac{n_i}{n_*} > 0,$$

that is

$$\sum_{i=1}^{r} 2^{q_i}(q_i - q) > 0.$$

Now note that the sum of the positive terms in $\sum_{i=1}^{n} 2^{q_i}(q_i - q)$ is at least 2^{q_1}. Furthermore, the absolute value of the sum of the negative terms is at most $2^{q_1-2} + 2(2^{q_1-3}) + 3(2^{q_1-4}) + \ldots + (q_1 - 1)$. So the proof is complete if one can show that,

$$2^{q_1} \geq 2^{q_1-2} + 2(2^{q_1-3}) + 3(2^{q_1-4}) + \ldots + (q_1 - 1).$$

The last inequality follows directly from the fact that $\sum_{i=1}^{q_1-1} i \cdot 2^{1-i} < \sum_{i=1}^{\infty} i \cdot 2^{1-i} = 4$ (Lemma 1.2). This completes the proof of the Lemma. $\Box$

Using this Lemma, it follows that

$$\chi(C) \geq \prod_{i=1}^{r} \chi(A_i) = 2^{\sum_{i=1}^{r} \frac{1}{2} n_i \log n_i - \sum_{i=1}^{r} n_i(1+o(1))}$$

$$> 2^{\sum_{i=1}^{r} \frac{1}{2} n_i \log(\sum_{i=1}^{r} n_i) - \sum_{i=1}^{r} n_i - n(1+o(1))}$$

$$= 2^{\frac{1}{2} n \log n - n - n(1+o(1))}$$

$$= 2^{\frac{1}{2} n \log n - n(2+o(1))}.$$

Thus we have a $(0,1)$ matrix C of order n, with $\chi(C)$ of the required order of magnitude. To obtain a $(-1,1)$ matrix, simply apply the map Φ described in subsection 2.2. Of course, the matrix $\Phi(C)$ has order $(n+1)$, but since we are dealing with asymptotic behaviour, this does not make any difference. This completes the proof of the Main Theorem. $\Box$

Remark Since $\det(S \diamond T) = \det S \det T$, and all the basic matrices of order 2^{n_i} we use have determinant -1 (see Remark at the end of subsection 2.2), the $(0,1)$ matrix C we just constructed has determinant of absolute value 1, and $|\det \Phi(C)| = 2^{n-1}$. This means that all the matrices constructed have minimum possible determinants.

3 Applications

3.1 Maximal norms of inverse matrices

In this subsection we estimate the maximum possible norms of inverses of $(0,1)$ and $(-1,1)$ matrices of order n. This is motivated by possible applications in numerical algebra. In particular, we answer the problem of Graham and Sloane mentioned in section 1. We also observe here that several quantities, including these norms, are closely related to the condition number of a matrix with $(0,1)$ or $(-1,1)$-entries.

Let B be a matrix of order n. The L_1, L_2, and *spectral* norms of B are defined as follows

$$\|B\|_1 = \max_i \sum_{j=1}^{n} |b_{ij}|, \|B\|_2 = \sqrt{\sum_{ij} b_{ij}^2}, \|B\|_s = \sup_{x \neq 0} \frac{|Bx|}{|x|}.$$

Let $\lambda_i(B)$ and $\sigma_i(B)$ be the eigenvalues and singular values of B in decreasing order of absolute value. Thus, $\sigma_i(B) = \sqrt{\lambda_i(B^t B)}$. The ratio $c(B) = \sigma_1(B)/\sigma_n(B)$ is an alternative formula for the condition number of B. It is useful to note that B and B^{-1} have the same condition number. The following properties are standard facts in linear algebra,

$$\sigma_n \leq |\lambda_n|, \ \|B\|_s = \sigma_1 \geq |\lambda_1| \text{ and } \|B\|_2^2 = \sum_{i=1}^{n} \sigma_i^2$$

Let $\mathcal{B}_n^i = \{A^{-1} | A \in \mathcal{A}_n^i, A \text{ invertible}\}$. Denote by $f_i(n)$, $e_i(n)$, $s_i(n)$ and $c_i(n)$ the following quantities: $\max_{B \in \mathcal{B}_n^i} \|B\|_1$, $\max_{B \in \mathcal{B}_n^i} \|B\|_2$, $\max_{B \in \mathcal{B}_n^i} \|B\|_s$ and $\max_{B \in \mathcal{B}_n^i} c(B)$, respectively. As shown below, all these quantities are closely related to the last one which is the maximum possible condition number of a matrix in $\mathcal{A}_n^i$. Moreover, $e_1^2(n) = \mu(n)$, where μ is defined in section 1.

Theorem 3.1.1. *For $i = 1, 2$, $f_i(n), e_i(n), s_i(n), c_i(n)$ have order of magnitude $2^{\frac{1}{2}n\log n-\theta(n)}$. More precisely, each of these quantities can be lower-bounded by $2^{(1/2)n\log n-n(2+o(1))}$, and upper-bounded by $2^{\frac{1}{2}n\log n-n(1+o(1))}$.*

Proof. By the definitions, and the above properties, $\|B\|_1$, $\|B\|_2$ and $\|B\|_s$ satisfy:

$$\chi(B^{-1}) \le \|B\|_i \le n\chi(B^{-1})$$

for $i = 1, 2$, and

$$n^{-1/2}\|B\|_2 \le \sigma_1 = \|B\|_s \le \|B\|_2.$$

Thus

$$n^{-1/2}\chi(B^{-1}) \le \sigma_1 = \|B\|_s \le n\chi(B^{-1})$$

The estimate concerning the L_1, L_2 and spectral norms follow immediately from the Main Theorem by taking the maxima in the inequalities above over the sets $B \in \mathcal{B}_n^i$ for $i = 1, 2$.

To estimate $c(n)$, first note that $\sigma_n(B) = \sigma_1(B^{-1})$. Moreover, $\sigma_1(B^{-1}) \le \|B^{-1}\|_2 \le n$, and $\sigma_1(B^{-1}) \ge |\lambda_1(B^{-1})| \ge |\det B^{-1}|^{1/n} \ge 1$. Thus, $1/n \le \sigma_n(B) \le 1$. This implies that

$$n^{-1/2}\chi(B^{-1}) \le c(B) \le n^2\chi(B^{-1}).$$

Again by maximizing over the sets $\mathcal{B}_n^i$, we deduce the desired estimate for $c_i(n)$ from the Main Theorem. $\square$

3.2 Flat simplices

In this subsection we estimate the minimum possible distance between a vertex and the opposite facet in a nontrivial simplex determined by $n+1$ vertices $P_1, P_2, \ldots, P_{n+1}$ of the unit hypercube $\{0,1\}^n$. Let $d(P_i)$ denote the distance from P_i to the hyperplane spanned by the other n points. The quantity we are interested in is $d(n) = \min_{P_1,P_2,..,P_{n+1}} \min_i d(P_i)$, where the minimum is taken over all indices i, and all possible configurations $P_1, P_2, \ldots, P_{n+1}$.

Without loss of generality, one can suppose that in the optimum configuration $P_{n+1} = \mathbf{0}$ and $d(n) = d(P_{n+1})$. Thus, the problem of determining $d(n)$ is equivalent to the problem of determining the minimum distance from the origin to a hyperplane spanned by vertices of the unit hypercube that does not go through the origin.

Let P be the $(0,1)$ matrix of order n whose rows are the points P_i. The distance from the origin to the hyperplane H spanned by the points P_i is

$$d(\mathbf{0}, H) = (\sum_{i=1}^{n}(\sum_{j=1}^{n} u_{ij})^2))^{-1/2}$$

as shown, for example, in [5], where u_{ij} are the entries of P^{-1}.

The following bounds for $d(n)$ are proved in [10], where the lower bound follows from Hadamard Inequality, and the upper bound is established by an appropriate construction.
Proposition 3.2.1 [10] *$d(n)$ satisfies the following inequalities:*

$$1.618^{-n} \geq d(n) \geq \frac{1}{2n^{3/2}}(\frac{4}{n})^{n/2}.$$

The lower bound is asymptotically $2^{-\frac{1}{2}n\log n+n(1+o(1))}$. Here we prove that $d(n)$ is upper-bounded by $\chi_1^{-1}(n)$, thus determining the asymptotic behaviour of $d(n)$.

Theorem 3.2.2 *$d(n)$ satisfies:*

$$2^{-\frac{1}{2}n\log n+n(1+o(1))} \leq d(n) \leq \chi_1^{-1}(n) \leq 2^{-\frac{1}{2}n\log n+n(2+o(1))}.$$

Proof. We construct the required simplex explicitly. It suffices to show that for every matrix $C \in \mathcal{A}_n^1$ one can construct a simplex for which the distance between a vertex and the opposite facet is at most $\chi(C)^{-1}$, since one can, in particular, take the matrix $C \in \mathcal{A}_n^1$ constructed in the proof of the Main Theorem. Given C, let v_i be the point represented by the i^{th} column vector of C. By reordering the rows and columns we can assume that $|\det C_{11}/\det C| = \chi(C) \geq 2^{(1/2)n\log n-n(2+o(1))}$. Let us denote by v the vertex $(1,0,0,..,0)$ of the hypercube. It is well known that $|\det C| = n!$ VolV_1, where V_1 is the simplex spanned by $\mathbf{0}$ and $v_1, v_2, .., v_n$. Similarly, $|\det C_{11}| = n!$Vol V_2, where V_2 is the simplex spanned by 0, v, and $v_2, .., v_n$. Denote by H the hyperplane through $\mathbf{0}$ and $v_2, v_3, ..., v_n$. Then

$$\chi(C)^{-1} = \frac{|\det C|}{|\det C_{11}|} = \frac{\text{Vol } V_2}{\text{Vol} V_1} = \frac{\text{dist}(v_1, H)}{\text{dist}(v, H)}$$

However, $\text{dist}(v, H) \leq \text{dist}(v, \mathbf{0}) = 1$. This implies that $\text{dist}(v_1, H) \leq \chi(C)^{-1}$, completing the proof.□

Remark. If $n = 2^m - 1$, by subsection 2.2, there are matrices C for which C^{-1} has a column in which every element is large, that is, $|\det C_{1i}/\det C| \geq 2^{\frac{1}{2}n\log n-n(2+o(1))}$ for every $1 \leq i \leq n$. This means that the above argument applies for all v_i. In geometric terms, it means that every vertex of V_1 except $\mathbf{0}$ is very close to the opposite facet.

In order to find a hyperplane close to the origin, one can choose an element of the automorphism group Aut$\{0,1\}^n$ which maps v_1 to $\mathbf{0}$. Then the images of the other n points of V_1 span a hyperplane determined by vertices of $\{0,1\}^n$, which is of distance $d(v_1, H)$ from the origin. In terms of the matrix C, this can be described in the following way. Starting with the matrix C in the proof, proceed as follows.

• Extend C to an $(n+1) \times n$ matrix C_1 by adding the zero vector $\mathbf{0}$ as the last row.

• Subtract the first row v_1 from each row of C_1 to get a matrix whose first row is $\mathbf{0}$, and whose remaining rows form an $n \times n$ matrix C_2.

• In C_2 replace all -1 entries by 1 entries, thus getting a $(0,1)$ matrix. The row vectors of this matrix span a hyperplane with distance $d(v_1, H)$ from the origin.

The problem of finding a flat simplex in the unit hypercube $(0,1)^n$ and that of finding a flat simplex in the hypercube $\{-1,1\}^n$ are the same, up to a factor of 2. But the hyperplane problem is different, since the origin is not a vertex of $\{-1,1\}^n$. However, the latter problem may also be solved easily, using the geometric interpretation of the map Φ, described in the previous section. If the vertices P_i of $(0,1)^{n-1}$ span a hyperplane H_1 with distance d from the origin in $\mathbf{R}^{n-1}$, then the vertices Q_i of $\{-1,1\}^n$, defined as in section 2 by Φ, span a hyperplane H_2 with distance less than d from the origin, since all P_i are contained in H_2.

3.3 Threshold gates with large weights

A *threshold gate* of n inputs is a function $F : \{-1,1\}^n \mapsto \{-1,1\}$ defined by

$$F(x_1, \ldots, x_n) = sign(\sum_{i=1}^{n} w_i x_i - t),$$

where $w_1, \ldots, w_n, t$ are reals called *weights*, chosen in such a way that the sum $\sum_{i=1}^{n} w_i x_i - t$ is never zero for $(x_1, \ldots, x_n) \in \{-1,1\}^n$. Threshold gates are the basic building blocks of Neural Networks, and have been studied extensively. See, e.g., [12] and its references. It is easy to see that every threshold gate can be realized with integer weights, and it is interesting to know how large these weights must be, in the worst case.

Let us call a threshold gate $F : \{-1,1\}^n \longrightarrow \{-1,1\}$ as above *recognizable*, and let us say that it is *recognized* by the pair (w,t). Given such a function F, there are many pairs (w,t) one can use to recognize F, and we are interested in the pair with minimum weight vector w, i.e., with weight vector of minimum possible l_∞ norm. We denote by $w(F)$ the l_∞ norm of this vector. (Note that the weight t can always be chosen to be at most $||w||_1 \leq n||w||_\infty$, and hence $w(F)$ supplies a bound for all weights.)

Let $\mathcal{F}_n$ be the set of all recognizable functions on $\{-1,1\}^n$. Define $w(n) = \max_{F \in \mathcal{F}} w(F)$. Our purpose is to describe the asymptotic behaviour of $w(n)$.

It has been proved by many researchers that if F is recognizable, then it can be recognized by integer weights satisfying $|w_i| \leq 2^{-n}(n+1)^{(n+1)/2} = 2^{\frac{1}{2}n\log n - n(1+o(1))}$. (See, e.g., [15].) Therefore, $w(n) \leq 2^{\frac{1}{2}n\log n - n(1+o(1))}$.

Håstad [11] proved that this upper-bound is nearly sharp for the case $n = 2^m$, by constructing a recognizable function which requires weights as large as $(1/2n)e^{-4n^\beta} 2^{\frac{1}{2}n\log n - n}$, where $\beta = \log(3/2) < 1$. We have exploited some of his ideas in the construction of ill-conditioned matrices in subsection 2.1.

However, if n is not a power of 2, no construction which requires weights close to the upper-bound is known. Of course, as suggested in [11], one may consider n_0, the largest power of 2 that does not exceed n, and use the construction for this number. This implies that $w(n) \geq w(n_0) = 2^{\frac{1}{2}n_0\log n_0 - n_0(1+o(1))}$. However, for n close to $2n_0$, this only gives $w(n) \geq 2^{\frac{1}{4}n\log n - n(1/2+o(1))}$, which is roughly the square root of the upper-bound.

As an application of the Main Theorem we construct here, for every n, a recognizable function F, which requires weights of absolute value at least $2^{\frac{1}{2}n\log n - n(2+o(1))}$. This determines the asymptotic behaviour of $w(n)$ up to an exponential factor.

Theorem 3.3.1 *$w(n)$ has order of magnitude $2^{\frac{1}{2}n\log n - \theta(n)}$. More precisely,*

$$2^{\frac{1}{2}n\log n - n(2+o(1))} \leq w(n) \leq 2^{\frac{1}{2}n\log n - n(1+o(1))}.$$

Proof. We have to prove the lower-bound. To this end, we construct an explicit function which requires such large weights.

Consider an ill-conditioned $(-1,1)$ matrix C of order n constructed in the Main Theorem, where $\chi(C) \geq 2^{\frac{1}{2}n\log n - n(2+o(1))}$. For convenience, suppose $\chi(C) = |\det C_{11}/\det C|$.

Let $v_1, v_2, \ldots, v_n$ be the row vectors of C. Define F on the v_i in the following way: $F(v_i) = sign(-1)^{i+1}\det C_{i1}\det C$ if $\det C_{1i} \neq 0$, otherwise $F(v_i) = 1$.

Since F is defined on n independent vectors, one can extend F to a recognizable odd function as follows. Choose a hyperplane H through the origin such that

• H does not contain any vertex of the cube $\{-1, 1\}^n$.

• All the points v_i, where $F(v_i) = 1$ are on one side of H, and all the points with $F(v_i) = -1$ are on the other side.

Since the hyperplane spanned by the v_i does not contain the origin, it is clear that such an H exists. Therefore, there is a weight vector w' such that $F(v_i) = sign < v_i, w' >$. Now extend F to all the vertices of the cube by defining $F(v) = sign < v, w' >$ for all v. Since w' is not orthogonal to any vertex vector of the cube, $F(v)$ is either -1 or 1, and hence F is recognizable by the pair $(w', 0)$. We next show that $w(F)$ satisfies the required lower-bound.

Let (w, t) be any integral pair that recognizes F. Since F is odd, $sign(< v, w > -t = -sign(< -v, w > -t)$ for all $(-1, 1)$ vector v. Hence $| < v, w > | > |t|$ for all v. This means that the pair $(w, 0)$ also recognizes F. Thus we may and will assume that $t = 0$.

Consider the vector $a = Cw$. Since w is integral, so is a. By the definition of F, it follows that $sign(a_i) = F(v_i)$. Now consider the equalities above as a system of linear equations with the variables w_i. By Cramer's rule we have

$$w_1 = \frac{\det C_1}{\det C} = \sum_{i=1}^{r} (-1)^{i+1} \frac{a_i \det C_{i1}}{\det C}$$

where C_1 is the matrix obtained from C by replacing its first column by a. By the definition of $F(v_i)$, all the terms in the right hand side are non-negative. Hence w_1 is at least as large as the first term:

$$w_1 \geq a_1 \frac{\det C_{11}}{\det C} \geq \chi(C),$$

since $|a_1| \geq 1$. This completes the proof.□

Remark. If n is a power of 2, a slightly better bound can be given, using the estimate in subsection 2.1. This special case is essentially the result of Hastad [11], with a somewhat different proof.

3.4 Coin weighing

Coin-weighing problems deal with the determination or estimation of the minimum possible number of weighings in a regular balance beam that enable one to find the required information about the weights of the coins. These questions have been among the most popular puzzles during the last fifty years, see, e.g., [9] and its many references. Here we study the following variant of the old questions, which we call the *all equal problem.*

Given a set of m coins, we wish to decide if all of them have the same weight or not, when various conditions about the weights are known in advance.

The case of *generic* weights, considered in [14], will be of special interest. In this case we assume that for the set $\{w_1, w_2, \ldots, w_t\}$ of possible weights of a coin, there is no set of integers $\lambda_1, \ldots, \lambda_n$ not all zero satisfying $\sum_{i=1}^t \lambda_i = \sum_{i=1}^t \lambda_i w_i = 0$. This assumption is motivated by the the fact that if we assume that the differences between the weights, which are supposed to be equal, are caused by effects of many independent sources, we should not expect any algebraic relation between the possible weights. In addition, the definition of generic weights is general enough to contain the basic case of two arbitrary distinct weights; every set $\{w_1, w_2\}$, $(w_1 \neq w_2)$ is generic.

Let $m(n)$ denote the maximum possible number of coins of generic potential weights for which the above problem can be solved in n weighings. It is not difficult to check (see [13], [14]) that $m(n) \geq 2^n$. To see this, note that trivially $m(1) = 2$, and that if we already know some m coins that have the same weight, then we can, in one additional weighing, compare them to m new coins and either conclude that not all coins have the same weight, in case the weighing is not balanced, or conclude that all $2m$ coins have the same weight, in case the last weighing is balanced. Hence $m(n+1) \geq 2m(n)$ for every n, implying that $m(n) \geq 2^n$.

Somewhat surprisingly, this is not tight. In [14] it is shown that $m(n) > 4.18^n$ and that $m(n) \leq \frac{3^n-1}{2}(n+1)^{(n+1)/2}$. A more general (though less explicit) bound for $m(n)$ is given in the following Theorem proved in [14].

Theorem 3.4.1. *Define $\gamma(n) = \max\{g(B),\ B \in \mathcal{B}\}$, where $g(B)$ denotes the minimum l_1 norm of a non-trivial integral solution of $Bx = 0$, and where $\mathcal{B}$ denotes the set of all $n \times n+1$ $(-1, 0, 1)$ matrices of rank n. Then*

$$\frac{3^n - 1}{2}\, \gamma(n) \geq m(n) \geq \gamma(n).$$

For a matrix $B \in \mathcal{B}$, it is easy to see that the vector $b = ((-1)^{i+1} \det B_i)_{i=1}^{n+1}$, where B_i is the square matrix obtained from B by deleting the i^{th} column, satisfies $Bb = 0$. Since B has rank n, every solution of $Bx = 0$ is a multiple of b. Hence

$$g(B) = \frac{\sum_{i=1}^{n+1} |\det B_i|}{gcd\{|\det B_i|\}_{i=1}^{n+1}}$$

where gcd stands for greatest common divisor. The main result of this subsection presented in Theorem 3.4.2 below, applies the above theorem together with our Main Theorem and improves the lower-bound of $m(n)$ up to only an exponential factor apart from the upper-bound. We also slightly improve the upper-bound by a factor of roughly $e^{1/2}$.

Theorem 3.4.2. $\frac{3^n-1}{2}(n+1)n^{(n-1)/2} \geq m(n) \geq 2^{\frac{1}{2}n\log n - n(2+o(1))}$.

Proof. To prove the upper-bound, it suffices to show that $\gamma(n) \leq (n+1)n^{(n-1)/2}$. Consider an $n \times (n+1)$ matrix B with entries $0, -1, 1$. If there are at least two rows of B that contain no zero entries, then each submatrix B_i contains at least two rows with $\{-1, 1\}$ entries. Adding one of them to the other, we get a matrix with a row all of whose entries are $0, 2$ or -2, and thus its determinant is divisible by 2. Hence all the numbers $|\det B_j|$ are divisible by 2. Thus, in this case $g(B) \leq \sum_{j=1}^{n+1} |\det B_j|/2$.

By adding to B a row $(b_1, \ldots, b_{n+1})$ of $\{-1, 1\}$ entries, where $b_j = sign(B_j)$, we obtain a matrix B' satisfying $|det(B')| = \sum_{j=1}^{n+1} |det(B_j)|$. By Hadamard Inequality, $|det(B')| \leq (n+1)^{(n+1)/2}$ and hence in this case

$$g(B) \leq \frac{(n+1)^{(n+1)/2}}{2} < (n+1)n^{(n-1)/2},$$

as needed.

It remains to bound $g(B)$ in case each of the rows of B, but possibly one, contains at least one zero. In this case, by Hadamard Inequality and with B' as above,

$$g(B) \le \sum_{j=1}^{n+1} |det(B_j)| = |det(B')| \le (n+1)n^{(n-1)/2}.$$

Since B was arbitrary, the desired result follows.

In order to prove the lower-bound, we construct, for every n, a $(0,1)$ and a $(-1,1)$ matrix of size $n \times n+1$, the γ of which is at least the claimed lower-bound. In fact, our construction has an even stronger property, which is described in the next Proposition. We note that both constructions, that of a $(0,1)$ matrix as well as that of a $(-1,1)$ matrix will be applied later, and we thus describe both although any one of them suffices to prove Theorem 3.4.2.

To state the proposition, we need some new notation. Let B be an $n \times n+1$ matrix of rank n, and let x be a non-trivial vector satisfying $Bx = 0$. Define $\xi(B) = \max_{1 \le i,j \le n+1, x_j \ne 0} |x_i/x_j|$. Note that ξ is well defined and is independent of the choice of x, since B has rank n. In fact, by a standard fact from linear algebra (mentioned above) the vector $((-1)^{i+1} \det B_i)_{i=1}^{n+1}$, where B_i is the square matrix obtained from B by deleting its i^{th} column, is a solution of the equation $Bx = 0$. Thus,

$$\xi(B) = \max_{1 \le i,j \le n+1, \det B_j \ne 0} |\det B_i / \det B_j|.$$

Proposition 3.4.3. *For every n, there is a $(0,1)$ $n \times (n+1)$ matrix B of rank n such that $\xi(B) \ge 2^{\frac{1}{2}n\log n - n(2+o(1))}$. There is also a $(-1,1)$ matrix with the same property.*

Proposition 3.4.3 easily supplies the lower bound in Theorem 3.4.2, since $\gamma(B)$ is at least $\xi(B)$. This follows from the following observation. If x is a non-trivial integral vector such that $Bx = 0$, and $\xi(B) = |x_p/x_q|$, then $\sum_{i=1}^{n+1} |x_i| \ge x_p \ge \xi(B)|x_q| \ge \xi(B)$.

Proof of Proposition 3.4.3.

The $(0,1)$ case. Pick a $(0,1)$ ill-conditioned matrix C of order n, such that $\chi(C) = |\det C_{11}/\det C| \ge 2^{\frac{1}{2}n\log n - n(2+o(1))}$. The matrix B is obtained from C by adding to its right a column $a = (1,0,0,\dots,0)$. Thus B has size $n \times (n+1)$ and rank n. Moreover,

$$\xi(B) \ge |\frac{\det B_1}{\det B_{n+1}}| = |\frac{\sum_{i=1}^{n}(-1)^{n+i}a_i \det C_{i1}}{\det C}|.$$

Observe that $a_1 = 1$ and $a_i = 0$ for all $i > 1$, implying that

$$\xi(B) \ge |\det C_{11}/\det C| = \chi(C) \ge 2^{\frac{1}{2}n\log n - n(2+o(1))}.$$

The $(-1,1)$ case. Again consider an ill-conditioned $(-1,1)$ matrix C with the same property as above. The matrix B is obtained by adding to the right side of C a $(-1,1)$ vector a, which will be defined later. As before, we have:

$$\xi(B) \ge |\frac{\det B_1}{\det B_{n+1}}| = |\frac{\sum_{i=1}^{n}(-1)^{n+i}a_i \det C_{i1}}{\det C}|.$$

Choose $a_i \in \{-1, 1\}$ such that each term in the sum in the numerator is non-negative. Hence the numerator is at least $\det C_{11}$. Thus,

$$\xi(B) \geq |\det C_{11}/\det C| = \chi(C) = 2^{\frac{1}{2}n\log n - n(2+o(1))}.$$

This completes the proof of Proposition 3.4.3 and implies the assertion of Theorem 3.4.2 as well. □

Although the existence of a weighing process follows from the last proposition by Theorem 3.4.1, we describe it here, for the sake of completeness. Once a matrix B (either a $(0,1)$ matrix or a $(-1,1)$ matrix) with the property described in Proposition 3.4.3 is found, the weighing process for solving the all equal problem for at least $\xi(B)$ coins using n weighings is as follows:

Weighing process

• By changing the sign of some columns of B, if needed, we may assume that there is a nontrivial solution of $Bx = 0$ which is non-negative. Choose such a solution w with the minimum possible l_1 norm. (This can be found by taking the smallest integral multiple of the basic solution $(\det B_i)_{i=1}^{n+1}$ with an appropriate sign.) Consider a set Ω of $m = \sum_{i=1}^{n} w_i$ coins. Clearly, $m \geq \xi(B)$. Let u_i, $i = 1, 2, \ldots, n+1$ denote the columns of B.

• Let W be the matrix obtained from B by duplicating each column u_i w_i times. Thus W is an $n \times m$ matrix. Index the columns of W by the coins of Ω. Let r_i denote the i^{th} row of W, and let v_j denote its j^{th} column.

• To define the i^{th} weighing ($1 \leq i \leq n$), consider the i^{th} row r_i of W. Let L_i be the set of coins corresponding to 1 entries, and let R_i be the set of coins corresponding to -1 entries in r_i. In the i^{th} weighing, we compare the weights of these two sets of coins.

• If there is an unbalanced weighing, we conclude that the coins are not weight-uniform. If all weighings are balanced, we conclude that the coins are of the same weight.

The proof of the fact that this weighing process does solve the all equal problem for coins of generic weights is not difficult. Here we sketch it for the case of two distinct weights.

Proof. Since $Bw = 0$, the number of 1 entries and -1 entries in any row of W is equal, and thus if any weighing is unbalanced, we can conclude that there are unequal weights. Suppose now all weighings are balanced. Indirectly, assume the coins are not weight-uniform. Let Ω' be the set of lighter coins. Since all weighings are balanced, L_i and R_i must contain the same number of lighter coins for all i. This implies that $\sum_{k\in\Omega'} v_k = 0$. Since each v_k is one of the vectors u_i, $1 \leq i \leq n+1$, this yields $\sum_{i=1}^{n+1} w'_i u_i = 0$, where w'_i is the multiplicity of u_i in the (multi-) set $\{v_k, k \in \Omega'\}$. But the last equation is equivalent to $Bw' = 0$, where $w' = (w'_1, w'_2, \ldots, w'_{n+1})$. Moreover, since Ω' is a proper nonempty subset of Ω, w' is not zero and $\|w'\|_1 < \|w\|_1$, a contradiction.□

The proof for the general case of more than 2 potential generic weights is similar. Let Ω' be the set of coins of some fixed weight. By the generic assumption we still have $|\Omega' \cap L_i| = |\Omega' \cap R_i|$ for all i, and one can conclude the proof in the same way. On the other hand, without the generic assumption, the situation changes drastically. Here is a brief discussion of this case (for more details see [3], [4]).

Let $m(n,k)$ denote the maximum possible number m such that given a set of m coins out of a collection of coins of k unknown distinct weights, one can decide if all the coins have the same weight or not using n weighings in a regular balance beam. In particular, $m(n,2)$ corresponds to the generic case considered above, in the special case there are two weights. Surprisingly, it turns out that $m(n,k)$ for $k \geq 3$ is much smaller than $m(n,2)$ $(= n^{(\frac{1}{2}+o(1))n}.)$ In [3] it is proved that for every $3 \leq k \leq n+1$, $m(n,k) = \Theta(n \log n / \log k)$. This indicates that the generic assumption is crucial.

However, we can prove that in case there is no assumption about the weights of the coins, our weighing process still works properly if we are given only one distinguished coin known to be either the lightest or the heaviest one. Here is a description of this process.

Let $M(n)$ denote the maximum possible number m such that given a set of m coins out of a collection of coins of an arbitrary number of unknown distinct weights, and given a distinguished coin which is known to be either the heaviest or the lightest one among the given m coins, one can decide if all the coins have the same weight or not using n weighings in a regular balance beam. Note that the distinguished coin may be either the heaviest or the lightest, and it is not known in advance which of the two it is. If there are only two possible weights, then any coin is distinguished, and hence this is a generalization of the basic case of two potential weights.

Theorem 3.4.4. $M(n) \geq 2^{\frac{1}{2}n \log n - n(2+o(1))}$

Proof. Suppose that the distinguished coin has the smallest weight (the proof is the same for the other case). To prove the inequality we prove that in case the matrix B in the weighing process is constructed from an ill-conditioned $(0,1)$ matrix C then the process also applies in the present situation.

First note that when B is constructed from a $(0,1)$ matrix then the standard solution $(-1)^{i+1} \det B_i$ is minimal, since $|\det B_{n+1}| = |\det C| = 1$ (see the remark at the end of the proof of the Main Theorem). Thus, the last column of W has multiplicity 1. Associate this column with the distinguished coin, and the other columns with the remaining coins. We show that if all weighings are balanced, then all coins have the same weight. Let τ_i be the weight of the coin associated to the column v_i, and let τ be the vector with coordinates τ_i. Since all weighings are balanced $W\tau = 0$. In addition, $W\mathbf{1}_m = 0$. Thus $W(\tau - \tau_m \mathbf{1}_m) = 0$. Note that $\tau_m = \min \tau_i$, implying that the vector $\tau - \tau_m \mathbf{1}_m$ has non-negative coordinates and its last coordinate is zero. Thus the product $W(\tau - \tau_m \mathbf{1}_m)$ is a linear combination of the first n columns of B, with non-negative coefficients. Since these n columns are independent (in fact they are the columns of C), their linear combination is zero iff all the coefficients are zero. This implies that $\tau_i - \tau_m = 0$ for all i, i.e., all coins have the same weight. □

3.5 Indecomposable hypergraphs

A *multi-hypergraph* H on a set X of n vertices is a collection of (not necessarily distinct) subsets of X, called edges. The *degree* of a vertex i in X is the number of subsets in the collection containing it. A (not necessarily induced) *sub-hypergraph* of H is a sub (multi)-set of H. A hypergraph is *regular* if all its vertices have the same degree. Let $D(n)$ be the maximum degree d so that there exists a regular hypergraph H with degree d, containing no proper nontrivial regular sub-hypergraph. We call such a hypergraph H *indecomposable.* The problem of estimating the value of $D(n)$ is motivated by some questions in Game Theory and was considered by various researchers (see [8] and its references). Huckeman and Jurkat proved that $D(n)$ is finite, (this was reproved by Alon and Berman, [1], using a different approach). The best known upper bound for $D(n)$ was given by Huckeman, Jurkat and Shapley (see [8])

$$D(n) \leq (n+1)^{(n+1)/2}.$$

In the other direction, Shapley showed that $D(n) > 2^{n-1}/(n-1)$ for every $n > 2$. This was improved by van Lint and Pollak, who showed that for all $n > 2$

$$D(n) \geq 2^{n-3} + 1.$$

Here we improve this lower-bound by showing that $D(n) \geq 2^{\frac{1}{2}n \log n - n(2+o(1))}$. This determines the asymptotic behaviour of $D(n)$ showing that it is $n^{(\frac{1}{2}+o(1))n}$.

Theorem 3.5.1 *$D(n)$ has order of magnitude $2^{\frac{1}{2}n \log n - O(n)}$. More precisely,*

$$2^{\frac{1}{2}n \log n + o(n)} \geq D(n) \geq 2^{\frac{1}{2}n \log n - n(2+o(1))}.$$

Proof. The upper-bound follows from the result of Huckeman, Jurkat and Shapley mentioned above. We thus have to prove the lower bound.

Consider a $(0,1)$, $n \times n$ matrix D and a non-negative integral vector $w = (w_1, \ldots, w_n)$. A multi-hypergraph $H = H(D,w)$ is defined by D and w as follows. The vertex-set of H is $\{1, 2, \ldots, n\}$. The edge-set consists of w_j copies of the set $\{i | d_{ij} = 1\}$, for every $j \leq n$. Therefore, there are n multi-edges. In other words, H is the multi-hypergraph with D as vertex-edge incidence matrix and the j^{th} edge has multiplicity w_j.

Now suppose D is a non-singular $(0,1)$ matrix of order n, for which the unique vector x such that $Dx = \mathbf{1}_n$ is non-negative. Let $N(D)$ be the minimal positive integer such that $w_i = N(D)x_i$ is integer for every index i. It is easy to verify that, in this case, the multi-hypergraph $H = H(D,w)$ is regular of degree $N(D)$. Furthermore, by the minimality of N, H is indecomposable. To estimate $N(D)$, note that $Nx_j \geq 1 \geq x_i$, for every x_i and $x_j \neq 0$, hence $N \geq \max_{i,j,\ x_j \neq 0} x_i/x_j$.

In order to prove the Theorem, we construct a non-singular $n \times n$ matrix D such that the unique solution of $Dx = \mathbf{1}_n$ is non-negative, and $N(D)$ is large.

Consider an $n \times (n+1)$ $(-1,1)$ matrix B, with the property described in proposition 3.4.3. Let w be a non-trivial vector satisfying $Bw = 0$. By reordering the columns, we can assume that $\xi(B)) = |w_1/w_2|$

By changing the sign of some columns of B, if needed, one can assume that w is non-negative. Moreover, by changing the sign of some rows, we can also assume that the last column is $-\mathbf{1}_n$. Let u_i denote the i^{th} column vector. The equality $Bw = 0$ implies that

$$\begin{aligned}
&\textstyle\sum_{i=1}^{n+1} w_i u_i = 0 \\
\iff &\textstyle\sum_{i=1}^{n} w_i u_i = w_{n+1}\mathbf{1}_n \\
\iff &\textstyle\sum_{i=1}^{n} \frac{w_i}{w_{n+1}} u_i = \mathbf{1}_n \\
\iff &\textstyle\sum_{i=1}^{n} \frac{w_i}{w_{n+1}} (u_i + \mathbf{1}_n) = (1 + \sum_{i=1}^{n} \frac{w_i}{w_{n+1}})\mathbf{1}_n \\
\iff &\textstyle\sum_{i=1}^{n} 2\frac{w_i}{w_{n+1}} (1 + \sum_{i=1}^{n} \frac{w_i}{w_{n+1}})^{-1} v_i = \mathbf{1}_n
\end{aligned}$$

where $v_i = \frac{1}{2}(u_i + \mathbf{1}_n)$. Note that the v_i are $(0,1)$ vectors. Let D be the $n \times n$ matrix with v_i as column vectors. We next prove that D satisfies the required properties.

1. D is non-singular. Suppose there is a non-trivial linear relation $\sum_{i=1}^n y_i v_i = 0$. In terms of u_i this means that $\sum_{i=1}^n y_i(u_i + \mathbf{1}_n) = 0$, or equivalently that $\sum_{i=1}^n y_i u_i + \sum_{i=1}^n y_i \mathbf{1}_n = 0$. The last equation means that the vector $(y_1, y_2, \ldots, y_n, -\sum_{i=1}^n y_i)$ is a solution of the system $Bx = 0$, which is a contradiction, since every solution of this system is either non-negative or non-positive. Thus D is non-singular.

2. The solution of $Dx = \mathbf{1}_n$ is $x = (2\frac{w_i}{w_{n+1}}(1 + \sum_{i=1}^n \frac{w_i}{w_{n+1}})^{-1})_{i=1}^n$. It is clear that x is non-negative. Furthermore,

$$N \geq \max_{1 \leq i,j \leq n, x_j \neq 0} |x_i/x_j| = \max_{1 \leq i,j \leq n, w_j \neq 0} \frac{2\frac{w_i}{w_{n+1}}(1 + \sum_{i=1}^n \frac{w_i}{w_{n+1}})^{-1}}{2\frac{w_j}{w_{n+1}}(1 + \sum_{i=1}^n \frac{w_i}{w_{n+1}})^{-1}}$$

$$= \max_{1 \leq i,j \leq n, w_j \neq 0} w_i/w_j = w_1/w_2 = \xi(B)$$

Thus $N(D) \geq \xi(B) \geq 2^{\frac{1}{2}n \log n - n(2+o(1))}$. This completes the proof. $\square$

4 Concluding remarks

• In case n is a power of 2, all the bounds using ill-conditioned matrices in our theorems can be improved, using Theorem 2.1.1, which gives a slightly better bound than the Main Theorem.

• Although the function $m(n, 2)$ is monotone by definition, it is not clear that so is the following version of its inverse. For an integer m, let $n(m)$ denote the minimum integer n such that given a set of m coins out of a collection of coins of two unknown distinct weights, one can decide if all the coins have the same weight or not using n weighings in a regular balance beam. It is not clear if for $m' < m$ the inequality $n(m') \leq n(m)$ holds, since the existence of an efficient weighing algorithm for m does not seem to imply the existence of an efficient one for a smaller number of coins. Using our techniques here we can, however, determine the asymptotic behaviour of $n(m)$ and show that

$$n(m) = (2 + o(1))\frac{\log m}{\log \log m},$$

where the $o(1)$-term tends to zero as m tends to infinity. A similar remark holds for the more general case of generic weights.

• In subsection 3.5 we prove that for all n, there is a $(0,1)$ matrix D of order n such that $N(D) \geq 2^{\frac{1}{2}n \log n - n(2+o(1))}$. Here, too, considering an appropriate inverse function is of interest. For every positive integer m, let $t(m)$ be the smallest number such that there is an invertible $(0,1)$ matrix D of order $t(m)$, for which the equation $Dx = \mathbf{1}_{t(m)}$ has a non-negative solution and $N(D) = m$. Our result implies that there are infinitely many values of m for which

$$t(m) \leq (2 + o(1))\frac{\log m}{\log \log m}.$$

It is not clear, however, that $t(m) \leq O(\log m)$ holds for all m. The estimate of $t(m)$ seems to be more difficult than that of $n(m)$. See [2] for some results on this question and on a related combinatorial problem.

• One can show that $M(n)$ is super-multiplicative by the following observation.

Put $m_1 = M(n_1)$, $m_2 = M(n_2)$. Given a collection of $m_1 m_2$ coins together with a distinguished one known to be either the heaviest or the lightest, we first apply the algorithm to the first m_1 coins (including the distinguished one), and use n_1 weighings to decide if all these coins have the same weight. If not, the algorithm ends. Otherwise, we split all coins into groups of size m_1, where the first group is the one consisting of the m_1 coins we already know to be equal. Viewing the groups as new coins, note that the first one must be either the heaviest or the lightest group. We can thus apply the algorithm and check the m_2 groups in n_2 weighings. If all the groups have the same weight, so do all the coins, and otherwise, not all coins are identical.

It is not clear if the function $m(n)$ corresponding to weighing coins with generic potential weights, the function $D(n)$ representing the maximum possible degree of indecomposable hypergraphs, or the function $w(n)$ describing the maximum required size of weights of threshold gates are super-multiplicative.

5 Acknowledgement

We would like to thank Imre Bárány, Anders Bjorner, László Lovász, Dmitry Kozlov, Imre Ruzsa and Günter Ziegler for many helpful discussions and comments.

References

[1] N. Alon and K. Berman, *Regular hypergraphs, Gordon's lemma, Steinitz's lemma and Invariant theory*, J. Combinatorial Theory, Ser. A 43 (1986), 91-97.

[2] N. Alon, D. J. Kleitman, K. Pomerance, M. Saks and P. D. Seymour, *The smallest $n-$uniform hypergraph with positive discrepancy*, Combinatorica 7 (1987), 151-160.

[3] N. Alon and D. N. Kozlov, *Coins with arbitrary weights*, J. Algorithms, to appear.

[4] N. Alon, D.N. Kozlov and V. H. Vũ, *The geometry of coin-weighing problems*, Proc. 37^{th} IEEE FOCS, IEEE (1996), 524-532.

[5] R. J. T. Bell, *An elementary treatise on coordinate geometry of three dimensions,* Second Edition, Macmillan, London, 1931.

[6] J. H. E. Cohn, *On the value of determinants,* Proc. Amer. Math. Soc 14 (1963), 581-588.

[7] G. H. Golub and J. H. Wilkinson, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Rev. 18 (1976), 578-619.

[8] J. E. Graver, *A survey of the maximum depth problem for indecomposable exact covers,* in: *Infinite and finite sets*, Colloq. Math. Soc. János Bolyai 10, North Holland, Amsterdam (1973), pp. 731-743.

[9] R. K. Guy and R. J. Nowakowsky, *Coin-weighing problems*, Amer. Math. Monthly 102 (1995), 164-167.

[10] R. L. Graham and N. J. A. Sloane, *Anti-Hadamard matrices*, Linear Algebra and its Applications 62 (1984), 113-137.

[11] J. Håstad, *On the size of weights for threshold gates,* SIAM J. Discrete Math. 7 (1994), 484-492.

[12] J. Hertz, R. Krogh and A. Palmer, *An Introduction to the Theory of Neural Computation*, Addison-Wesley, Reading, MA 1991.

[13] F. K. Hwang, P. D. Cheng and X. D. Hu, *A new competitive algorithm for the counterfeit coin problem*, Infor. Proc. Letters 51 (1994), 213-218.

[14] D. N Kozlov and V. H. Vũ , *Coins and cones*, J. Combinatorial Theory, Ser. A 78 (1997), 1-14.

[15] S. Muroga, *Threshold Logic and its Application,* Wiley-Interscience, New York, 1971.

[16] J. H. Wilkinson, *Note on matrices with a very ill-conditioned eigenproblem*, Numer. Math. 19 (1972), 176-178.