Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Higher degree Erdős distinct evaluations problem

(Extended abstract)

Simone Costa* Stefano Della Fiore† Andrea Ferraguti*

Abstract

Let $\Sigma = \{a_1, ..., a_n\}$ be a set of positive integers with $a_1 < \cdots < a_n$ such that all 2^n subset sums are distinct. A famous conjecture by Erdős states that $a_n > c \cdot 2^n$ for some constant c, while the best result known to date is of the form $a_n > c \cdot 2^n/\sqrt{n}$.

In this paper, we propose a generalization of the Erdős distinct sum problem that is in the same spirit as those of the Davenport and the Erdős-Ginzburg-Ziv constants recently introduced in [7] and in [6]. More precisely, we require that the non-zero evaluations of the m-th degree symmetric polynomial are all distinct over the subsequences of Σ. Even though these evaluations can not be seen as the values assumed by the sum of independent random variables, surprisingly, the variance method works to provide a nontrivial lower bound on a_n. Indeed, the main result here presented is to show that

$$a_n > c_m \cdot 2^{\frac{n}{m}}/n^{1-\frac{1}{2m}}.$$

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-043`

1 Introduction

For any $n \geq 1$, consider sets $\{a_1, ..., a_n\}$ of positive integers with $a_1 < \cdots < a_n$ whose subset sums are all distinct. A famous conjecture, due to Paul Erdős, is that $a_n \geq c \cdot 2^n$

*DICATAM - Sez. Matematica, University of Brescia, Via Branze 38, I-25123 Brescia, Italy. E-mails: simone.costa@unibs.it, andrea.ferraguti@unibs.it.

†DI, University of Salerno, Via Giovanni Paolo II 132, Fisciano, Italy. E-mail: s.dellafiore@unisa.it.

for some constant $c > 0$. Using the variance method, Erdős and Moser [10] (see also [1] and [13]) were able to prove that $a_n \geq 1/4 \cdot n^{-1/2} \cdot 2^n$. No advances have been made so far in removing the term $n^{-1/2}$ from this lower bound, but there have been several improvements on the constant factor, including the work of Dubroff, Fox, and Xu [11], Guy [12], Elkies [9], Bae [4], and Aliev [3]. In particular, the best currently known lower bound states that $a_n \geq (1+o(1))\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{n}}2^n$. Two simple proofs of this result, first obtained unpublished by Elkies and Gleason, are presented in [11]. In the other direction, the best-known construction is due to Bohman [5] (see also [14]), who showed that there exist arbitrarily large such sets with $a_n \leq 0.22002 \cdot 2^n$.

Several variations on the problem appear during the years such as [2] and [8]. In this paper, we generalize the Erdős distinct sum problem by requiring that the non-zero evaluations of the m-th degree symmetric polynomial are all distinct over the sub-sequences of Σ. The problem here considered is inspired by those of the Davenport and the Erdős-Ginzburg-Ziv constants recently introduced in [7] and in [6].

More formally, given a sequence of real numbers $\Sigma = \{a_1, \dots, a_n\}$ and a subset $A \subseteq [1, n]$, we define the m-th (degree) evaluatio $e^m_\Sigma(A) = \sum_{\substack{\{i_1,\dots,i_m\}\subseteq A \\ i_1<\cdots<i_m}} a_{i_1}\cdots a_{i_m}$, where we adopt the convention that $e^m_\Sigma(A) = 0$ if $|A| < m$.

Problem 1.1. *For every positive integer n, find the least positive $M = M(n)$ such that there exists an increasing sequence $\Sigma = (a_1, \dots, a_n)$ of real numbers with $a_i \in [0, M]$ for every i such that for all distinct $A_1, A_2 \subseteq [1, n]$ of size at least m we have that $|e^m_\Sigma(A_1) - e^m_\Sigma(A_2)| \geq 1$.*

A sequence as in Problem 1.1 will be called *M-bounded m-th evaluation distinct.*

In Section 2, we provide lower bounds on the values of M in Problem 1.1 using the variance method proving that

$$M > c_m \cdot 2^{\frac{n}{m}}/n^{1-\frac{1}{2m}}.$$

Then, in Section 3, we derive an upper bound presenting a direct construction.

2 Lower Bounds

One first lower bound to the value of M of Problem 1.1 can be provided using the pigeonhole principle. Indeed, since the number of non-zero evaluations of e^m_Σ is $2^n - \sum_{i=0}^{m-1}\binom{n}{i} = (1+o(1))2^n$, these evaluations are spaced at least by one, and each of these is smaller than $e^m_\Sigma([1, n]) \leq \binom{n}{m}M^m \leq n^m M^m/c_m$, it follows that $M > c_m \cdot 2^{\frac{n}{m}}/n$.

Now we see that using the variance method (see [1], [10] or [12]), it is possible to improve this lower bound.

Theorem 2.1. *Let $\Sigma = (a_1, \dots, a_n)$ be an m-th evaluation distinct sequence in $\mathbb{R}$ (resp. $\mathbb{Z}$) that is M-bounded. Then*

$$M > (1+o(1))\frac{2^{1-\frac{1}{m}}((m-1)!)^{\frac{1}{m}}}{3^{\frac{1}{2m}}}\frac{2^{\frac{n}{m}}}{n^{1-\frac{1}{2m}}}.$$

Proof. Let $\Sigma = (a_1, \ldots, a_n)$ be such a sequence of real (resp. integer) numbers. Pick a subset A uniformly at random from $2^{[1,n]}$ and define the real random variable $X = e^m_\Sigma(A)$. We denote by $\mu := \mathbb{E}[X]$ and $\sigma^2 := \mathbb{E}[X^2] - \mu^2$ respectively the expected value and the variance of the random variable X. Clearly, $\mu = 1/2^n \sum_{A\subseteq[1,n]:\ |A|\geq m} e^m_\Sigma(A)$. Here we have that the monomial $a_{i_1} \ldots a_{i_m}$ appears in the evaluation $e^m_\Sigma(A)$ whenever A contains $i_1, \ldots, i_m$ which happens for 2^{n-m} subsets of $[1,n]$. Therefore, we have that $\mu = e^m_\Sigma([1,n])/2^m$. By definition of variance we have that:

$$2^n\sigma^2 = \sum_{A\subseteq[1,n]} (e^m_\Sigma(A) - \mu)^2 = \sum_{A\subseteq[1,n]} \left(\sum_{\substack{i_1<i_2<\cdots<i_m \\ i_1,\ldots,i_m\in A}} a_{i_1}\ldots a_{i_m} - \sum_{\substack{i_1<i_2<\cdots<i_m \\ i_1,\ldots,i_m\in[1,n]}} \frac{a_{i_1}\ldots a_{i_m}}{2^m} \right)^2 .$$

Due to the symmetry of e^m_Σ, there exist coefficients $C_1, \ldots, C_m$ such that the latter sum can be written as follows:

$$C_0 \sum_{\substack{i_1<i_2<\ldots<i_{2m} \\ i_1,\ldots,i_{2m}\in[1,n]}} a_{i_1}\ldots a_{i_{2m}} + C_1 \sum_{\substack{i_1<i_2<\ldots<i_{2m-1} \\ i_1,\ldots,i_{2m-1}\in[1,n]}} \sum_{\ell\in[1,2m-1]} a_{i_1}a_{i_2}\ldots a^2_{i_\ell}\ldots a_{i_{2m-1}} + \tag{1}$$
$$+ \ldots + C_m \sum_{\substack{i_1<i_2<\ldots<i_m \\ i_1,\ldots,i_m\in[1,n]}} a^2_{i_1}\ldots a^2_{i_m}.$$

One can prove that $C_0 = 0$, $C_1 = 2^{n-2m}\binom{2m-2}{m-1}$ and $C_k = O(2^n)$ for every $k \in \{2, \ldots, m\}$. This can be seen since the coefficient of $a_{i_1}\ldots a_{i_{2m}}$ is $\binom{2m}{m}$ times that obtained by taking the term $a_{i_1}\ldots a_{i_m}$ from the first $(e^m_\Sigma(A) - \mu)$ in the product and $a_{i_{m+1}}\ldots a_{i_{2m}}$ from the second one. Then, the coefficient of $a^2_{i_1}\ldots a_{i_2}\ldots a_{i_{2m-1}}$ is $\binom{2m-2}{m-1}$ times that obtained taking the term $a_{i_1}\ldots a_{i_m}$ from the first $(e^m_\Sigma(A) - \mu)$ in the product and $a_{i_1}a_{i_{m+1}}\ldots a_{i_{2m-1}}$ from the second one. Symmetrically, the same is true for every term $a_{i_1}\ldots a^2_{i_\ell}\ldots a_{i_{2m-1}}$. Finally, the coefficient of $a^2_{i_1}\ldots a^2_{i_k}a_{i_{k+1}}\ldots a_{i_{2m-k}}$ is $\binom{2m-2k}{m-k}$ times that obtained taking the term $a_{i_1}\ldots a_{i_m}$ from the first $(e^m_\Sigma(A) - \mu)$ in the product and $a_{i_1}\ldots a_{i_k}a_{i_{m+1}}\ldots a_{i_{2m-k}}$ from the second one. Summing up, we can rewrite equation (1) as

$$2^n\sigma^2 = C_1 \sum_{\substack{i_1<i_2<\ldots<i_{2m-1} \\ i_1,\ldots,i_{2m-1}\in[1,n]}} \sum_{\ell\in[1,2m-1]} a_{i_1}a_{i_2}\ldots a^2_{i_\ell}\ldots a_{i_{2m-1}} + \tag{2}$$
$$+O(2^n)\left(\sum_{k=2}^{m} C_k \sum_{\substack{i_1<i_2<\ldots<i_{2m-k} \\ i_1,\ldots,i_{2m-k}\in[1,n]}} \sum_{\substack{\ell_1<\ldots<\ell_k \\ \ell_1,\ldots,\ell_k\in[1,2m-k]}} a_{i_1}a_{i_2}\ldots a^2_{i_{\ell_1}}\ldots a^2_{i_{\ell_k}}\ldots a_{i_{2m-k}} \right).$$

In equation (2), each C_k multiplies a sum of $\binom{n}{2m-k}\cdot\binom{2m-k}{k} < \frac{n^{2m-k}}{(2m-2k)!k!}$ terms. Since

a_n is the largest element of the sequence, we get:

$$2^n\sigma^2 < \frac{n^{2m-1}}{(2m-2)!}\binom{2m-2}{m-1}2^{n-2m}a_n^{2m}(1+o(1)) = \left(\frac{n^{2m-1}}{((m-1)!)^2}2^{n-2m}a_n^{2m}\right)(1+o(1)). \tag{3}$$

On the other hand, for $|A| \geq m$, the evaluations $e_\Sigma^m(A)$ are all different and spaced at least by one, and hence we have that $(e_\Sigma^m(A)-\mu)^2$ assumes at least $\frac{1}{2}(2^n - \sum_{i=0}^{m-1}\binom{n}{i})$ different values. Since the sum $\sum_{A\subseteq[1,n]}(e_\Sigma^m(A)-\mu)^2$ is minimized when the values are around μ and are spaced by one, we obtain the lower bound:

$$\frac{1+o(1)}{12}2^{3n} = 2\sum_{i=0}^{\frac{1}{2}(2^n-\sum_{i=0}^{m-1}\binom{n}{i})} i^2 \leq 2^n\sigma^2. \tag{4}$$

To conclude the proof, it is enough to compare (3) and (4). □

3 Upper bounds

In this section we provide an upper bound to the value of M in Problem 1.1 by presenting the following direct construction.

Lemma 3.1. *Let ϵ_1, ϵ_2 be two reals such that $\epsilon_1 > \epsilon_2 > 0$ and let $m \geq 2$ be an integer. Then for every n large enough the sequence $\Sigma = (a_1, a_2, \ldots, a_n)$, where $a_i = (2+\epsilon_1)^n - (2+\epsilon_2)^{i-1}$ for $i = 1, 2, \ldots, n$, is m-evaluation distinct.*

Proof. Suppose by contradiction there exists two distinct subsets $B, C \subseteq [1,n]$ such that

$$|e_\Sigma^m(B) - e_\Sigma^m(C)| < 1\,. \tag{5}$$

For an arbitrary subset $S \subseteq [1,n]$ with $|S| \geq m$, by definition we have:

$$e_\Sigma^m(S) = \sum_{j=0}^{m}(-1)^j(2+\epsilon_1)^{(m-j)n}\binom{|S|-j}{m-j}\sum_{\substack{\{i_1,i_2,\ldots,i_j\}\subseteq S\\ i_1<i_2<\ldots<i_j}}(2+\epsilon_2)^{i_1+i_2+\ldots+i_j-j}\,. \tag{6}$$

We first show that inequality (5) implies $|B| = |C|$. Suppose without loss of generality that $|B| > |C|$. Then (6) implies that:

$$e_\Sigma^m(B) - e_\Sigma^m(C) = (2+\epsilon_1)^{mn}\left[\binom{|B|}{m} - \binom{|C|}{m}\right] + \sum_{j=1}^{m}(-1)^j(2+\epsilon_1)^{(m-j)n}$$
$$\left[\binom{|B|-j}{m-j}\sum_{\substack{\{i_1,i_2,\ldots,i_j\}\subseteq B\\ i_1<i_2<\ldots<i_j}}(2+\epsilon_2)^{i_1+i_2+\ldots+i_j-j} - \binom{|C|-j}{m-j}\sum_{\substack{\{i_1,i_2,\ldots,i_j\}\subseteq C\\ i_1<i_2<\ldots<i_j}}(2+\epsilon_2)^{i_1+i_2+\ldots+i_j-j}\right]. \tag{7}$$

Now it can be seen that each term in the first summation of equation (7) is of order $O\left(n^m(2+\epsilon_1)^{mn}\left(\frac{2+\epsilon_2}{2+\epsilon_1}\right)^{jn}\right)$, for $j = 1, 2, \ldots, m$ and $n \to \infty$. Hence, asymptotically in n, we can rewrite (7) as $e^m_\Sigma(B) - e^m_\Sigma(C) = (2+\epsilon_1)^{mn}\left[\binom{|B|}{m} - \binom{|C|}{m}\right](1+o(1))$, since $\epsilon_1 > \epsilon_2$.This clearly contradicts (5), and hence we must have $|B| = |C|$.

Next, let t be an integer such that $|B| = |C| = t$ and let $B := \{b_1, b_2, \ldots, b_t\}$ and $C := \{c_1, c_2, \ldots, c_t\}$, where $b_1 < b_2 < \ldots < b_t$ and $c_1 < c_2 < \ldots < c_t$. Since $B \neq C$, there exists an integer $\ell \in [1, t]$ such that $b_\ell \neq c_\ell$ while $b_{\ell+1} = c_{\ell+1}$, $b_{\ell+2} = c_{\ell+2}, \ldots$, $b_t = c_t$. Suppose without loss of generality that $b_\ell > c_\ell$. Then we have:

$$|e^m_\Sigma(B) - e^m_\Sigma(C)| = \Bigg|(2+\epsilon_1)^{(m-1)n}\binom{t-1}{m-1}\left(\sum_{i\leq\ell}(2+\epsilon_2)^{b_i-1} - (2+\epsilon_2)^{c_i-1}\right) + \sum_{j=2}^{m}(-1)^{j-1}$$
$$(2+\epsilon_1)^{(m-j)n}\binom{t-j}{m-j}\left(\sum_{\substack{1\leq i_1<i_2<\ldots<i_j\leq t\\ i_1\leq\ell}}(2+\epsilon_2)^{b_{i_1}+b_{i_2}+\ldots+b_{i_j}-j} - (2+\epsilon_2)^{c_{i_1}+c_{i_2}+\ldots+c_{i_j}-j}\right)\Bigg|. \tag{8}$$

To conclude the proof, we need to lower bound equation (8). The summation formula for the geometric series implies that: $\sum_{i\leq\ell}(2+\epsilon_2)^{c_i-1} \leq \sum_{1\leq i\leq c_\ell}(2+\epsilon_2)^{i-1} < (2+\epsilon_2)^{c_\ell}/(1+\epsilon_2) \leq (2+\epsilon_2)^{b_\ell-1}/(1+\epsilon_2)$, and since each term in the summation over j in equation (8) is, as $n \to \infty$, of order $O\left(n^m(2+\epsilon_1)^{(m-1)n}(2+\epsilon_2)^{b_\ell-1}\left(\frac{2+\epsilon_2}{2+\epsilon_1}\right)^{(j-1)n}\right)$, we obtain the following lower bound:

$$|e^m_\Sigma(B) - e^m_\Sigma(C)| > \left|(2+\epsilon_1)^{(m-1)n}\binom{t-1}{m-1}(2+\epsilon_2)^{b_\ell-1}\left(1 - \frac{1}{1+\epsilon_2}\right)\right|(1+o(1)).$$

The theorem now follows since the right hand side of the above inequality is greater than 1 for sufficiently large n's. □

Along the same lines of Lemma 3.1, we can prove the following corollary. We do not report here the proof due to space limitations.

Corollary 3.2. *Let ϵ_1, ϵ_2 be two reals such that $\epsilon_1 > \epsilon_2 > 0$ and let $m \geq 2$ be an integer. Then for every n large enough the sequence $\Sigma = (a_1, a_2, \ldots, a_n)$, where $a_i = \lfloor(2+\epsilon_1)^n - (2+\epsilon_2)^{i-1}\rfloor$ for $i = 1, 2, \ldots, n$, is m-evaluation distinct.*

We observe that Corollary 3.2 holds also for $m = 1$ but we obtain a bound that is worse than the ones given in [5] and [14]. As an easy consequence of Corollary 3.2, one can prove the following theorem.

Theorem 3.3. *There exists a sequence $\Sigma = (a_1, a_2, \ldots, a_n)$ of n integers that is m-evaluation distinct and M-bounded such that $M \leq 2^{n+o(n)}$, for $n \to \infty$.*

References

[1] N. Alon and J. H. Spencer, The probabilistic method, 4th ed. Wiley, Hoboken, NJ, 2016.

[2] M. Axenovich, Y. Caro, R. Yuster, Sum-distinguishing number of sparse hypergraphs, European Journal of Combinatorics 112 (2023), 103712 .

[3] I. Aliev, Siegel's lemma and sum-distinct sets, Discrete Comput. Geom. 39 (2008), 59-66.

[4] J. Bae, On subset-sum-distinct sequences. Analytic number theory, Vol. 1, Progr. Math., 138, Birkhauser, Boston, 1996, 31-37.

[5] T. Bohman, A construction for sets of integers with distinct subset sums, Electron. J. Combin. 5 (1998), Research Paper 3, 14 pages.

[6] Y. Caro and J.R. Schmitt. Higher degree Erdős-Ginzburg-Ziv constants, Integers 22 (2022).

[7] Y. Caro, B. Girard and J.R. Schmitt. Higher degree Davenport constants over finite commutative rings, Integers 21 (2021).

[8] S. Costa, S. Della Fiore, M. Dalai: Variation on the Erdős distinct-sums problem. Discrete Applied Mathematics, 325 (2023), 172–185.

[9] N. D. Elkies, An improved lower bound on the greatest element of a sum-distinct set of fixed order, J. Combin. Theory Ser. A 41 (1986), 89-94.

[10] P. Erdős, Problems and results in additive number theory, Colloque sur la Theorie des Nombres, Bruxelles, 1955, 127-137.

[11] Quentin Dubroff, Jacob Fox and Max Wenqiang Xu, A note on the Erdős distinct subset sums problem, SIAM J. Discret. Math. 35 (2021), 322-324.

[12] R. K. Guy, Sets of integers whose subsets have distinct sums, Theory and practice of combinatorics, 141–154, North-Holland Math. Stud., 60, Ann. Discrete Math., 12, North-Holland, Amsterdam, (1982).

[13] R. K. Guy, Unsolved Problems in Intuitive Mathematics, Vol. I, Number Theory, Problem C8, Springer-Verlag (1981).

[14] W.F. Lunnon, Integer sets with distinct subset sums, Math. Compute, 50 (1988) 297-320.