Variations on the Erdős Distinct-Sums Problem

Stefano Della Fiore
joint work with Simone Costa and Marco Dalai

Università degli studi di Brescia – DICATAM
Università degli studi di Brescia – DII

October 23, 2025

Overview

1. Formulation of the original problem by Erdős
2. Some known results from literature
3. Variations on the original problem
4. Information-theoretic Interpretation
5. Our results

Formulation of the original problem

Let $\{a_1, ..., a_n\}$ be a set of positive integers with $a_1 < \ldots < a_n$ such that all 2^n subset sums are distinct.

Conjecture.

A famous conjecture by Erdős states that $a_n > c \cdot 2^n$ for some constant c.

Theorem.

The best results known to date are of the form $a_n > c \cdot 2^n/\sqrt{n}$ for some constant c.

Improving the factor $\sqrt{n}$ is a very hard task and so only the constant c has been improved in the past 70 years.

Known bounds from literature

Lower bounds on a_n

- Trivial one $\to a_n \geq \frac{1}{n}2^n$
- Erdős and Moser (1955) $\to a_n \geq \frac{1}{4} \cdot \frac{2^n}{\sqrt{n}}$
- Alon and Spencer (1981) $\to a_n \geq (1 + o(1))\frac{2}{3\sqrt{3}} \cdot \frac{2^n}{\sqrt{n}}$
- Elkies (1986) $\to a_n \geq (1 + o(1))\frac{1}{\sqrt{\pi}} \cdot \frac{2^n}{\sqrt{n}}$
- Guy (1981) $\to a_n \geq (1 + o(1))\frac{1}{\sqrt{3}} \cdot \frac{2^n}{\sqrt{n}}$
- Dubroff, Fox and Xu (2020) $\to a_n \geq (1 + o(1))\sqrt{\frac{2}{\pi}} \cdot \frac{2^n}{\sqrt{n}}$

Upper bounds on a_n

- Trivial one $\to a_n \leq 2^{n-1}$ (take each $a_i = 2^{i-1}$)
- Bohman (1998) $\to a_n \leq 0.22002 \cdot 2^n$

Trivial bound - Pigeonhole principle

The maximum possible sum we can get is

$$\begin{aligned}\sum_{i=1}^{n} a_i &\leq a_n + \overbrace{(a_n - 1)}^{a_{n-1}} + \overbrace{(a_n - 2)}^{a_{n-2}} + \ldots + \overbrace{(a_n - n + 1)}^{a_1} \\ &= n a_n - \frac{n(n-1)}{2} \,.\end{aligned}$$

Since in the interval $[1, \sum_i a_i]$ there must be at least $2^n - 1$ integers, by the Pigeonhole principle we have that

$$a_n \geq \frac{2^n}{n} \,.$$

Second moment method - Alon and Spencer

Let us fix a set $\{a_1, a_2, \ldots, a_n\}$ with distinct partial sums.

Let $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ be i.i.d random variables distributed as

$$\Pr(\epsilon_i = 1) = \Pr(\epsilon_i = -1) = 1/2 \text{ for all } i \in [1, n]\,.$$

Set $X = \sum_{i=1}^{n} \epsilon_i a_i$. Clearly X is symmetric around 0 and uniformly distributed over its support of size $2^n \to \mathbb{E}[X] = 0$.

Let us bound

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}\left[\left(\sum_i \epsilon_i a_i\right)^2\right] = \sum_i \mathbb{E}\left[\epsilon_i^2\right] a_i^2 + 2\sum_{i<j} \mathbb{E}[\epsilon_i \epsilon_j] a_i a_j \\ &= \sum_i a_i^2 \leq n a_n^2\,.\end{aligned}$$

Second moment method - Alon and Spencer - 2

By Chebyschev's inequality for any $\lambda > 1$ we have that

$$\Pr(|X| < \lambda\sqrt{n}a_n) > 1 - \frac{1}{\lambda^2}$$

But X is a discrete random variable with probabilities equal to 2^{-n} or zero in an interval and each possible outcome has the same parity. Therefore

$$\Pr(|X| < \lambda\sqrt{n}a_n) \leq (\lambda\sqrt{n}a_n + 1)2^{-n}$$

Thus

$$a_n \geq \frac{2^n(1-\lambda^{-2}) - 1}{\lambda\sqrt{n}}$$

Optimizing for $\lambda > 1$ we get $\lambda = \sqrt{3}$. Hence $a_n \geq (1 + o(1))\frac{2^n}{3\sqrt{3}\sqrt{n}}$

As before, let $X = \sum_{i=1}^{n} \epsilon_i a_i$, where ϵ_i is the same random variable defined before.

We have that $\mathbb{E}[X] = 0$ and $\mathrm{Var}[X] \leq n a_n^2$, then we can lower bound $\mathrm{Var}[X]$ as follows:

$$\begin{aligned}\mathrm{Var}[X] &\geq \frac{1}{2^n}\left(1^2 + (-1)^2 + 3^2 + (-3)^2 + \ldots + (2^n - 1)^2 + (1 - 2^n)^2\right) \\ &= \frac{2}{2^n}\sum_{i=1}^{2^{n-1}} (2i-1)^2 = \frac{4^n - 1}{3}\,.\end{aligned}$$

Hence

$$a_n \geq (1 + o(1))\frac{1}{\sqrt{3}}\frac{2^n}{\sqrt{n}}\,.$$

Let us consider the following generating function

$$G(z) := \prod_{i=1}^{n} \left(1 + \frac{1}{2} \left(z^{a_i} + z^{-a_i} \right) \right) ,$$

where the $\{a_1, \ldots, a_n\}$ satisfy the sum distinct property.

By the sum distinct property the constant term in the power series expansion (Laurent expansion) of $G(z)$ around 0 must be equal to 1. Therefore using the Cauchy's integral formula we have that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{it}) dt = 1 .$$

Let us consider the following generating function

$$G(z) := \prod_{i=1}^{n} \left(1 + \frac{1}{2} \left(z^{a_i} + z^{-a_i} \right) \right) ,$$

where the $\{a_1, \ldots, a_n\}$ satisfy the sum distinct property.

By the sum distinct property the constant term in the power series expansion (Laurent expansion) of $G(z)$ around 0 must be equal to 1. Therefore using the Cauchy's integral formula we have that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{it}) dt = 1 .$$

Elkies's method - 2

$$2\pi = \int_{-\pi}^{\pi} \left[\prod_{i=1}^{n} (1 + \cos(a_i t)) \right] dt = 2^n \int_{-\pi}^{\pi} \prod_{i=1}^{n} \cos^2(a_i t/2) dt$$

since $1 + \cos(2t) = 2\cos^2(t)$

$$> 2^n \int_{-\pi/a_n}^{\pi/a_n} \prod_{i=1}^{n} \cos^2(a_i t/2) dt > 2^n \int_{-\pi/a_n}^{\pi/a_n} \prod_{i=1}^{n} \cos^2(a_n t/2) dt$$

$$= 2^n \int_{-\pi/a_n}^{\pi/a_n} \cos^{2n}(a_n t/2) dt = 2^n \frac{2}{a_n} \int_{-\pi/2}^{\pi/2} \cos^{2n}(u) du$$

$$= \frac{2^{n+1}}{a_n} 2^{-2n} \binom{2n}{n} \pi \,.$$

Hence

$$a_n \geq (1 + o(1)) \frac{1}{\sqrt{\pi}} \frac{2^n}{\sqrt{n}}$$

$$2\pi = \int_{-\pi}^{\pi} \left[\prod_{i=1}^{n} (1 + \cos(a_i t)) \right] dt = 2^n \int_{-\pi}^{\pi} \prod_{i=1}^{n} \cos^2(a_i t/2) dt$$

$$> 2^n \int_{-\pi/a_n}^{\pi/a_n} \prod_{i=1}^{n} \cos^2(a_i t/2) dt > 2^n \int_{-\pi/a_n}^{\pi/a_n} \prod_{i=1}^{n} \cos^2(a_n t/2) dt$$

since $\cos(a_i t/2) > \cos(a_n t/2)$ for $t \in [-\frac{\pi}{a_n}, \frac{\pi}{a_n}]$ and $i \in [1, n-1]$

$$= 2^n \int_{-\pi/a_n}^{\pi/a_n} \cos^{2n}(a_n t/2) dt = 2^n \frac{2}{a_n} \int_{-\pi/2}^{\pi/2} \cos^{2n}(u) du$$

$$= \frac{2^{n+1}}{a_n} 2^{-2n} \binom{2n}{n} \pi \,.$$

Hence

$$a_n \geq (1 + o(1)) \frac{1}{\sqrt{\pi}} \frac{2^n}{\sqrt{n}}$$

$$
\begin{aligned}
2\pi &= \int_{-\pi}^{\pi} \left[\prod_{i=1}^{n} (1 + \cos(a_i t)) \right] dt = 2^n \int_{-\pi}^{\pi} \prod_{i=1}^{n} \cos^2(a_i t/2) dt \\
&> 2^n \int_{-\pi/a_n}^{\pi/a_n} \prod_{i=1}^{n} \cos^2(a_i t/2) dt > 2^n \int_{-\pi/a_n}^{\pi/a_n} \prod_{i=1}^{n} \cos^2(a_n t/2) dt \\
&= 2^n \int_{-\pi/a_n}^{\pi/a_n} \cos^{2n}(a_n t/2) dt = 2^n \frac{2}{a_n} \int_{-\pi/2}^{\pi/2} \cos^{2n}(u) du \\
&= \frac{2^{n+1}}{a_n} 2^{-2n} \binom{2n}{n} \pi \,.
\end{aligned}
$$

Hence

$$
a_n \geq (1 + o(1)) \frac{1}{\sqrt{\pi}} \frac{2^n}{\sqrt{n}}
$$

Isoperimetric method - Dubroff, Fox and Xu

Let $\mathcal{F}$ be a family of subsets of $[1, n]$, and define the vertex boundary as

$$\partial\mathcal{F} = \{A \in 2^{[1,n]} \setminus \mathcal{F} : |A \Delta B| = 1 \text{ for some } B \in \mathcal{F}\}$$

Theorem (Harper vertex-isoperimetric)

If $\mathcal{F}$ is a family of subsets of $[1, n]$ and $|\mathcal{F}| = 2^{n-1}$, then

$$|\partial\mathcal{F}| \geq \binom{n}{\lfloor n/2 \rfloor} = (1 + o(1))\sqrt{\frac{2}{\pi}}\frac{2^n}{\sqrt{n}}$$

Isoperimetric method - Dubroff, Fox and Xu - 2

Let $a = (a_1, a_2, \ldots, a_n)$ be a sum-distinct sequence $a_1 < \ldots < a_n$, and note that $a \cdot \epsilon \neq 0$ for all $\epsilon \in \{-1/2, 1/2\}^n$.

Let $\mathcal{F}$ be the set of all 2^{n-1} points ϵ such that $a \cdot \epsilon < 0$.

Then, $\eta \in \partial\mathcal{F}$ iff $\eta - e_i \in \mathcal{F}$ for some basis vector e_i. In addition, each $\eta \in \partial\mathcal{F}$ satisfies $0 < a \cdot \eta < a_n$.

By Harper we have $|\partial\mathcal{F}| \geq \binom{n}{\lfloor n/2 \rfloor}$. Therefore there exists distinct $\eta_1, \eta_2 \in \partial\mathcal{F}$ such that $|a \cdot (\eta_1 - \eta_2)| \leq a_n / \binom{n}{\lfloor n/2 \rfloor}$.

By the sum distinct property we have $|a \cdot (\eta_1 - \eta_2)| \geq 1$ and hence $a_n \geq \binom{n}{\lfloor n/2 \rfloor} = (1 + o(1))\sqrt{\frac{2}{\pi}}\frac{2^n}{\sqrt{n}}$.

Isoperimetric method - Dubroff, Fox and Xu - 2

Let $a = (a_1, a_2, \ldots, a_n)$ be a sum-distinct sequence $a_1 < \ldots < a_n$, and note that $a \cdot \epsilon \neq 0$ for all $\epsilon \in \{-1/2, 1/2\}^n$.

Let $\mathcal{F}$ be the set of all 2^{n-1} points ϵ such that $a \cdot \epsilon < 0$.

Then, $\eta \in \partial\mathcal{F}$ iff $\eta - e_i \in \mathcal{F}$ for some basis vector e_i. In addition, each $\eta \in \partial\mathcal{F}$ satisfies $0 < a \cdot \eta < a_n$.

By Harper we have $|\partial\mathcal{F}| \geq \binom{n}{\lfloor n/2 \rfloor}$. Therefore there exists distinct $\eta_1, \eta_2 \in \partial\mathcal{F}$ such that $|a \cdot (\eta_1 - \eta_2)| \leq a_n / \binom{n}{\lfloor n/2 \rfloor}$.

By the sum distinct property we have $|a \cdot (\eta_1 - \eta_2)| \geq 1$ and hence $a_n \geq \binom{n}{\lfloor n/2 \rfloor} = (1 + o(1))\sqrt{\frac{2}{\pi}}\frac{2^n}{\sqrt{n}}$.

Isoperimetric method - Dubroff, Fox and Xu - 2

Let $a = (a_1, a_2, \ldots, a_n)$ be a sum-distinct sequence $a_1 < \ldots < a_n$, and note that $a \cdot \epsilon \neq 0$ for all $\epsilon \in \{-1/2, 1/2\}^n$.

Let $\mathcal{F}$ be the set of all 2^{n-1} points ϵ such that $a \cdot \epsilon < 0$.

Then, $\eta \in \partial\mathcal{F}$ iff $\eta - e_i \in \mathcal{F}$ for some basis vector e_i. In addition, each $\eta \in \partial\mathcal{F}$ satisfies $0 < a \cdot \eta < a_n$.

By Harper we have $|\partial\mathcal{F}| \geq \binom{n}{\lfloor n/2 \rfloor}$. Therefore there exists distinct $\eta_1, \eta_2 \in \partial\mathcal{F}$ such that $|a \cdot (\eta_1 - \eta_2)| \leq a_n / \binom{n}{\lfloor n/2 \rfloor}$.

By the sum distinct property we have $|a \cdot (\eta_1 - \eta_2)| \geq 1$ and hence $a_n \geq \binom{n}{\lfloor n/2 \rfloor} = (1 + o(1))\sqrt{\frac{2}{\pi}}\frac{2^n}{\sqrt{n}}$.

Isoperimetric method - Dubroff, Fox and Xu - 2

Let $a = (a_1, a_2, \ldots, a_n)$ be a sum-distinct sequence $a_1 < \ldots < a_n$, and note that $a \cdot \epsilon \neq 0$ for all $\epsilon \in \{-1/2, 1/2\}^n$.

Let $\mathcal{F}$ be the set of all 2^{n-1} points ϵ such that $a \cdot \epsilon < 0$.

Then, $\eta \in \partial\mathcal{F}$ iff $\eta - e_i \in \mathcal{F}$ for some basis vector e_i. In addition, each $\eta \in \partial\mathcal{F}$ satisfies $0 < a \cdot \eta < a_n$.

By Harper we have $|\partial\mathcal{F}| \geq \binom{n}{\lfloor n/2 \rfloor}$. Therefore there exists distinct $\eta_1, \eta_2 \in \partial\mathcal{F}$ such that $|a \cdot (\eta_1 - \eta_2)| \leq a_n / \binom{n}{\lfloor n/2 \rfloor}$.

By the sum distinct property we have $|a \cdot (\eta_1 - \eta_2)| \geq 1$ and hence $a_n \geq \binom{n}{\lfloor n/2 \rfloor} = (1 + o(1))\sqrt{\frac{2}{\pi}}\frac{2^n}{\sqrt{n}}$.

Variations on the original problem

First variation.
The distinct-sums condition is weakened by only requiring that the sums of up to λn elements of the set be distinct with $0 < \lambda < 1$.

Second variation.
The elements $a_i \in \mathbb{Z}^k$ for some $k \geq 1$.

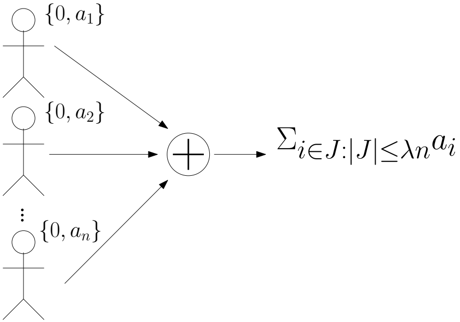
More formally...

Problem.
Let $\mathcal{F}_{\lambda,n}$ be the family of all subsets of $\{1,\ldots,n\}$ whose size is smaller than or equal to λn. We are interested in the minimum M such that there exists a sequence $\Sigma = (a_1,\ldots,a_n)$ in $\mathbb{Z}^k$, $a_i \in [0,M]^k\ \forall i$, such that for all distinct $A_1, A_2 \in \mathcal{F}_{\lambda,n}$, $S(A_1) \neq S(A_2)$, where

$$S(A) = \sum_{i\in A} a_i\,.$$

Remark.
For $\lambda = 1$ and $k = 1$ we obtain the same problem as the original one.

Signaling over a multiple access channel. Let $\{a_1, ..., a_n\}$ be a set, where $a_i \in \mathbb{Z}^k$ for some $k \geq 1$, such that the sums of up to λn elements of the set be distinct with $0 < \lambda < 1$. We have n trasmitters.



Each one can send a signal of amplitude a_i to the base station saying that it wants to start a communication session.

Proposition.

Let $\Sigma = (a_1, \ldots, a_n)$ be an $\mathcal{F}_{\lambda,n}$-sum distinct sequence in $\mathbb{Z}^k$ that is M-bounded. Then

$$M \geq \frac{1}{\lceil \lambda n \rceil} \cdot |\mathcal{F}_{\lambda,n}|^{1/k}.$$

Proof.

Since the maximum possible sum is at most $\lceil \lambda n \rceil M$ in each component, by the pigeonhole principle, we have that

$$M^k \geq \frac{1}{\lceil \lambda n \rceil^k} \cdot |\mathcal{F}_{\lambda,n}| .$$

□

Following the idea of Dubroff, Fox and Xu the previous bounds for $k = 1$ and $\lambda \geq 1/2$ can be improved as follow

Theorem

Let $\Sigma = (a_1, \ldots, a_n)$ be an $\mathcal{F}_{\lambda,n}$-sum distinct sequence in $\mathbb{Z}$ that is M-bounded. Then

$$M \geq (1 + o(1)) \cdot \begin{cases} \frac{1}{\sqrt{2\pi n}} \cdot 2^n & \text{if } \lambda = 1/2; \\ \sqrt{\frac{2}{\pi n}} \cdot 2^n & \text{if } \lambda \in]1/2, 1]. \end{cases}$$

Lower bound - Variance method for $k > 1$

Using the variance method, it is possible, to improve the previous bounds for $k > 1$.

Theorem

Let $\lambda \in [0, 1]$ and let $\Sigma = (a_1, \ldots, a_n)$ be an $\mathcal{F}_{\lambda,n}$-sum distinct sequence in $\mathbb{Z}^k$ that is M-bounded. Then

$$M \geq (1 + o(1)) \cdot \sqrt{\frac{4}{\pi n(k+2)}} \cdot \Gamma(k/2+1)^{1/k} \cdot |\mathcal{F}_{\lambda,n}|^{1/k}$$

where Γ is the gamma function.

The key step.

Consider a random variable $X = \sum_{i=1}^{n} \epsilon_i a_i$ where the random vectors $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ are uniformly distributed over the set $\mathcal{F}_{\lambda,n}$ with $\epsilon_i \in \{-1/2, 1/2\}$. It can be proved that for $\lambda \geq 1/2$

$$\mathbb{E}[\epsilon_i \epsilon_j] \leq 0$$

for each $i \neq j$. While for $\lambda < 1/2$ we have that

$$\mathbb{E}[\epsilon_i \epsilon_j] = O(1/n)\,.$$

Upper bound - Polynomial method

Using the polynomial method (Alon's combinatorial nullstellensatz) we get the following theorem.

Theorem

For any $\lambda < 1/3$, there exists a sequence $\Sigma = (a_1, \ldots, a_n)$ that is M-bounded positive integers and $\mathcal{F}_{\lambda,n}$-sum distinct with

$$M \geq \lambda^3 n^2 2^{f(\lambda)n},$$

where
$f(\lambda) = H(\lambda, \lambda, 1-2\lambda) = -2\lambda \log_2 \lambda - (1-2\lambda)\log_2(1-2\lambda)$.

Remark.

The previous bound is non-trivial for $\lambda < 3/25$.

Using the probabilistic method we get an improvement on the trivial bound (i.e., $c \cdot 2^{n/k}$) for $k > 1$ and $\lambda < 3/25$.

Theorem

Let

$$C_{\lambda,n} = \sqrt[k]{\frac{\lambda^2 n^2}{2\tau_\lambda} 2^{f(\lambda)\tau_\lambda}} \text{ and } \tau_\lambda = \left\lceil \frac{1}{\log_e 2 \cdot f(\lambda)} \right\rceil,$$

where $f(\lambda) = -2\lambda \log_2 \lambda - (1 - 2\lambda) \log_2 (1 - 2\lambda)$.

Then there exists a sequence $\Sigma = (a_1, \ldots, a_n)$ *of* $\left(C_{\lambda,n} \cdot 2^{f(\lambda)n/k}\right)$*-bounded elements of* $\mathbb{Z}^k$ *that is* $\mathcal{F}_{\lambda,n}$*-sum distinct.*

Improvements for $k = 1$ and $\lambda \in]3/25, 1/4]$

Using the Lunnon's construction it can be shown that if $\lambda < 1/4$ and n large enough, there exists a sequence $\Sigma = (a_1, \ldots, a_n)$ of M-bounded integers that is $\mathcal{F}_{\lambda,n}$-sum distinct with

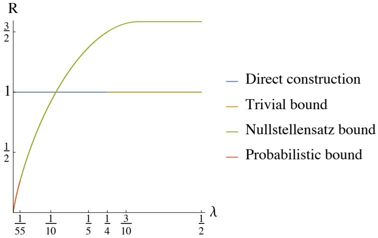
$$M = \frac{0,22096}{2} \cdot 2^n \,,$$

while for $\lambda < 1/8$ we get

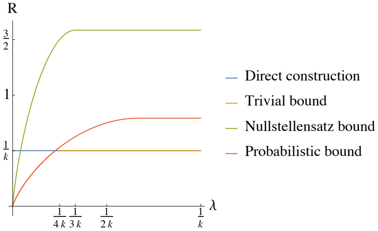
$$M = \frac{0,22096}{4} \cdot 2^n \,.$$

Remark.
For $\lambda < 1/4$ we can insert an additional a_i to the sequence found by Lunnon while for $\lambda < 1/8$ two elements can be added without violating the sum-distinct property.

Figure: Exponent of the upper bounds for $k = 1$ and for $k > 1$.



(a) $k = 1$



(b) $k > 1$

The problem over symmetric polynomials

Problem

For every positive integer n, find the least positive $M = M(n)$ such that there exists a sequence $\Sigma = (a_1, \ldots, a_n)$ of integers 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:

$$e^m_\Sigma(A_1) \neq e^m_\Sigma(A_2),$$

where

$$e^m_\Sigma(A) = \sum_{\substack{\{i_1,\ldots,i_m\}\subseteq A \\ i_1<\ldots<i_m}} a_{i_1} \cdots a_{i_m},$$

and we adopt the convention that $e^m_\Sigma(A) = 0$ if $|A| < m$.