Optimal Result for a Signed Sum Problem

Kecen Sha

January 15th, 2026

Abstract

This paper studies a signing problem from the 2021 China High School Mathematics League (Add-on Round). The main result of this work is a refined threshold: we define an explicit sequence c_n satisfying $c_n = \frac{n}{2} + O(1)$, depending on $n \bmod 8$, and prove that for **any** $n \in \mathbb{N}$, any subset $A \subseteq [n]$ with $|A| > c_n$ admits a signing with sum in $\{-1, 0, 1\}$ and c_n is optimal. Our proof improves upon a known result that established the same conclusion with $n \geq 564$. We improve the bound to $n \geq 38$ first. For small n, the result could be verified by a practical computational time, from which we confirm the optimality of c_n and thus provide an optimal result for this problem.

1 Background

The problem investigated in this paper originates from the 2021 China High School Mathematics League (Additional Round), a prestigious national mathematical competition. The original contest problem asked for the smallest constant $c > 0$ such that for every integer $n \geq 4$ and every subset $A \subseteq \{1, 2, \ldots, n\}$ with $|A| > cn$, there exists a signing function $f : A \to \{-1, 1\}$ satisfying

$$\left| \sum_{a \in A} f(a) \cdot a \right| \leq 1.$$

The official solution establishes that the optimal constant is $c = \frac{2}{3}$. The proof proceeds by a clever combinatorial argument combined with an auxiliary lemma on signed sums of bounded integers. Specifically, it considers the parity of $|A|$ and constructs a derived set $\bar{A}$ consisting of either consecutive differences $a_{i+1} - a_i$ (for even-sized A) or those differences together with the smallest element a_1 (for odd-sized A). Then it proves by induction that if $s(\bar{A}) = \sum\limits_{a \in \bar{A}} a \leq 2|\bar{A}|$, the signed sum can achieve 0 or 1. The fact $|A| > \frac{2}{3}n$ makes $\bar{A}$ satisfy our condition and thus the proof is finished.

Subsequently, Wang Bingyuan significantly refined this result by determining the exact threshold sequence c_n for all sufficiently large n (see in (1)) . He proved that for $n \geq 564$, the optimal value is given explicitly by c_n, which matches the asymptotic density $\frac{1}{2}$ and improves upon the competition's $\frac{2}{3}$ bound in the finite setting. However, his method relies on analytic estimates and structural lemmas that only become effective when n is large; it does not extend to smaller values of n , leaving a substantial gap for $n < 564$.

In this work, we completely resolve the problem for all positive integers n . We adopt a more refined and case-adaptive strategy: instead of a single universal criterion, we design multiple complementary screening algorithms tailored to different combinatorial structures of candidate sets. This multi-pronged approach, combining theoretical reductions with targeted computational verification, enables us to handle the remaining finite cases exhaustively. As a result, we establish the same explicit formula for c_n as Wang's for $n \geq 18$, and provide exact values for all smaller n , thereby achieving a full, non-asymptotic characterization of the optimal threshold sequence.

2 Main Results

Define $\{c_n\}_{n=1}^{+\infty}$ as follows: for $n \geq 18$, let

$$c_n = \begin{cases} 4k+1 & \text{if } n = 8k+2 \text{ or } 8k+3, \\ 4k+2 & \text{if } n = 8k+4, \ldots, 8k+7, \\ 4k+3 & \text{if } n = 8k+8 \text{ or } 8k+9. \end{cases} \tag{1}$$

For $n \leq 17$, let $\{c_i\}_{i=1}^{17} = \{0,1,2,2,3,4,4,5,5,5,6,6,7,7,7,8,8\}$.

The key result is stated below:

Theorem 1. *For any integer $n \geq 0$, and any subset $A \subseteq \{1,2,3,\ldots,n\}$, if $|A| > c_n$, then there exists a function $f : A \to \{1,-1\}$ satisfying*

$$\left| \sum_{a \in A} f(a) \cdot a \right| \leq 1. \tag{2}$$

Moreover, there exists $|A| = c_n$ such that we can not find f to let (2) hold.

A known result establishes Theorem 1 with $n \geq 564$. We first give a rigorous proof for $n \geq 38$ and then verify Theorem 1 by programming for small n.

Theorem 2 (Through theoretical derivation)**.** *For any integer $n \geq 38$, and any subset $A \subseteq \{1,2,3,\ldots,n\}$, if $|A| > c_n$, then there exists a function $f : A \to \{1,-1\}$ satisfying*

$$\left| \sum_{a \in A} f(a) \cdot a \right| \leq 1.$$

To prove Theorem 2, we first handle the case where there exists a subset $B \subset A$ such that either $B = \{1\}$ or $B = \{x, x+1\}$. In such cases, it suffices to consider $A \setminus B$. The remaining exceptional configurations can be dealt with manually.

Theorem 3. *For any integer $n \geq 38$, and any subset $A \subseteq \{1,2,3,\ldots,n\}$, if $|A| \geq \frac{n-5}{2}$, then there exists a function $f : A \to \{1,-1\}$ satisfying*

$$\left| \sum_{a \in A} f(a) \cdot a \right| \leq 2.$$

3 Notation and Definition

We will analyze two types of data structures: sets and multisets. A **set** contains only distinct elements, while a **multiset** permits repeated elements, each with a well-defined multiplicity. A set could be viewed as a multiset with each element counted with multiplicity 1. Without clarification, we assume all uppercase letters $A, B, C\cdots$ denote multisets. For any $r \geq 1$, denote multiset $I_r = \{1, 1\cdots, 1\}$ which contains r number of 1.

For multisets A and B, we write $A - B$ to mean $A \setminus (A \cap B)$. We use the standard notation $[n] = \{1, 2, \ldots, n\}$. For any multiset X, we adopt the following conventions:

1. $|X|$ denotes its cardinality, with multiplicities counted for multisets.
2. $X^{\uparrow} = \{x_i\}_{i=1}^{|X|}$ denotes the sequence of elements of X sorted in non-decreasing order;
3. $s(X) = \sum_{x\in X} x$ denotes the sum of elements, with multiplicities counted for multisets.
4. $\mathcal{L}(X, m) = |X \cap [m, +\infty)|$ denote the number of upper level set of X at level m.

For a set $A \subset \mathbb{N}^*$, if $|A|$ is even, define

$$\bar{A} = \{a_2 - a_1, a_4 - a_3, \ldots, a_{|A|} - a_{|A|-1}\}, \quad \tilde{A} = \{a_1, a_3 - a_2, \ldots, a_{|A|-1} - a_{|A|-2}\};$$

otherwise, define

$$\bar{A} = \{a_1, a_3 - a_2, \ldots, a_{|A|} - a_{|A|-1}\}, \quad \tilde{A} = \{a_2 - a_1, \ldots, a_{|A|-1} - a_{|A|-2}\}.$$

Next, we introduce key definitions relevant to our theorems and proofs.

Definition 1. *We say that A is r-**strongly expressive** if for every integer $0 \leq x \leq s(A \cup I_r)$ with $x \equiv s(A \cup I_r) \pmod 2$, there exists a function $f : A \cup I_r \to \{1, -1\}$ such that*

$$\sum_{a\in A\cup I_r} f(a)\cdot a = x.$$

Let S_r be the set of all sets that are r-strongly expressive.

From definition, $A \in S_0$ implies there exists a function $f : A \to \{1, -1\}$ such that

$$\left|\sum_{a\in A} f(a)\cdot a\right| \leq 1.$$

It is easy to see that if $A \in S_r$ and $s \leq r$, then $A \cup I_s \in S_{r-s}$.

Definition 2. *We say that a r-strongly expressive multiset A can **absorb** an integer $b \in \mathbb{N}$ if $A \cup \{b\}$ is still r-strongly expressive.*

We show a necessary and sufficient condition for when we can absorb an element.

Proposition 1. *Let A be a r-strongly expressive multiset and $b \in \mathbb{N}$, then A can absorb b if and only if $b \leq s(A) + r + 1$.*

Proof. Let $X = A \cup I_r$, then X can represent all values $s(X), s(X)-2, \dots, \varepsilon$, where $\varepsilon \in \{0, 1\}$. Adding b allows us to represent $b+\varepsilon, \dots, b+s(X)$, while adding $-b$ yields $b-\varepsilon, \dots$. If $b \le s(X)+1$, this sequence covers all residues modulo 2 down to 0 or 1, depending on $b+s(X) \pmod 2$. If $b \ge s(X)+2$, then any signed sum of $X \cup \{b\}$ cannot represent neither 1 nor 0. $\square$

Definition 3. *Let $A \subset \mathbb{N}^*$ be a set. Define its* ***residue energy*** *$E(A)$ as:*

$$E(A) = s(\bar{A}) + s(\tilde{A}) - |A|. \tag{3}$$

The residue energy reflects the extent of gap between consecutive elements. The smaller the residue energy, the smaller the typical difference $a_{i+1} - a_i$. This suggests A is dense and likely to be strongly expressive.

Proposition 2. *Let $A \subset [n]$, then $0 \le E(A) \le n - |A|$ and for any $m, l \ge 0$,*

$$E(A) \ge \mathcal{L}(\bar{A}, |\bar{A}|+1-m) \cdot (|\bar{A}|-m) + \mathcal{L}(\tilde{A}, |\tilde{A}|+1-l) \cdot (|\tilde{A}|-l). \tag{4}$$

Proof. Since $s(\bar{A}) \ge |\bar{A}|, s(\tilde{A}) \ge |\tilde{A}|, |\bar{A}| + |\tilde{A}| = |A|$, we have $E(A) \ge 0$. Let $A^{\uparrow} = \{a_i\}_{i=1}^{t}$, then

$$n \ge a_t = \sum_{i=0}^{t-1}(a_{i+1} - a_i) = \sum_{2 \nmid i-t}(a_{i+1} - a_i) + \sum_{2 \mid i-t}(a_{i+1} - a_i) = s(\bar{A}) + s(\tilde{A}).$$

So $E(A) \le n - |A|$. By definition of upper level set, we have

$$s(\bar{A}) \ge \mathcal{L}(\bar{A}, |\bar{A}|+1-m)(|\bar{A}|+1-m) + |\bar{A}| - \mathcal{L}(\bar{A}, |\bar{A}|+1-m);$$

$$s(\tilde{A}) \ge \mathcal{L}(\tilde{A}, |\tilde{A}|+1-l)(|\tilde{A}|+1-l) + |\tilde{A}| - \mathcal{L}(\tilde{A}, |\tilde{A}|+1-l),$$

adding them together we finish the proof. $\square$

4 Useful Lemmas

We now present a useful criterion for a strongly expressive multiset.

Lemma 1 (Sufficient condition for r-strong expressiveness)**.** *Let $A \subseteq \mathbb{N}^*$ and $r \in \mathbb{N}$. If for all $m \in [0, |A|) \cap \mathbb{N}$, we have*

$$\mathcal{L}(A, |A|+r+1-m) \le m,$$

then A is r-strongly expressive.

Proof. Let $A^{\uparrow} = \{a_1, \dots, a_s\}$. If $a_i \ge i+r+1$, then $\mathcal{L}(\bar{A}, i+r+1) > |\bar{A}| - i$, a contradiction. Hence $a_i \le i+r$ for all i. Consider the sequence $\{b_i\}_{i=1}^{r+s}$ such that $b_1 = \cdots = b_r = 1, b_{r+i} = a_i$, then $b_k \le k \le 1 + b_1 + \cdots b_{k-1}$ for all k. Hence, starting from $\varnothing$, the multiset can absorb b_i in a sequence. Thus, A is r-strongly expressive by definition. $\square$

The next lemma shows that if A is a subset of $[n]$ with $|A| \sim 0.4n$, then A is nearly 1-strongly expressive.

Lemma 2. *Let $A \subseteq [n]$ and $|A| > \frac{2(n+3)}{5}$. Then for all $m \in \{2, 3, \ldots, |\bar{A}| - 2\}$, we have*

$$\mathcal{L}(\bar{A}, |\bar{A}| + 2 - m) \leq m. \tag{5}$$

Proof. Let $t = |A|, s = |\bar{A}|$, if $\mathcal{L}(\bar{A}, |\bar{A}| + 2 - m) \geq m + 1$, by Proposition 2, we have

$$n - t \geq (m + 1) \cdot (s + 1 - m) \geq 3s - 3,$$

where the last inequality holds by the monotonicity of the quadratic in m. It contradicts that $s \geq t/2, t > \frac{2n+6}{5}$. □

5 Proof of Theorem 3

We argue by contradiction. Suppose A does not satisfy the condition, then any multiset generated by signed sum of the elements in A (with each used exactly once) is not 1-strongly expressive. For example, A itself and $\bar{A}$.

5.1 Case 1: $\mathcal{L}(\bar{A}, 3) < |\bar{A}|$.

Let $t = |A|$ and $s = \lceil t/2 \rceil = |\bar{A}|$. Since $t \geq \frac{n-5}{2} > \frac{2n+6}{5}$ for $n \geq 38$, Lemma 2 implies that $\mathcal{L}(\bar{A}, s + 1) \geq 1$. Consequently, there exists an index $j \equiv t \pmod 2$ such that

$$a_j - a_{j-1} \geq s + 1.$$

Without loss of generality, we assume $d = a_j - a_{j-1} = \max \bar{A}$.

Claim 1. $\mathcal{L}(\bar{A} - \{d\}, 2s + 1) = \mathcal{L}(\tilde{A}, 2s + 1) = 0$.

Proof. If not, $n - t \geq E(A) \geq 2s + d - 1 \geq 3s$, a contradiction to $t \geq \frac{n-5}{2}, n \geq 26$. □

Claim 2. *Let $A_1 = [A \cap (0, a_{j-1})] \cup \{0\}$ and $A_2 = A \cap (a_j, n]$. Then there exist $x, y \in A_1$ or $x, y \in A_2$ such that*

$$d - s \leq |x \pm y| \leq d + s.$$

Proof. Assume the contrary. Write $A_1^{\uparrow} = \{0 < x_1 < \cdots < x_p\}$ and $A_2^{\uparrow} = \{y_1 < \cdots < y_q\}$. Then $x_p \leq d - s - 1$; otherwise $x_p \geq d + s + 1$. But by Claim 1, we have $x_{i+1} - x_i \leq 2s + 1$ for all i, and the intermediate value principle yields a contradiction. Similarly, $y_q - y_1 \leq d - s - 1$.

Suppose that there are w number of $x_{i+1} - x_i, y_{i+1} - y_i$ (including x_1) being not 1. Then,

$$t - 3 + w = p + q + w - 1 \leq x_p + (y_q - y_1) \leq 2(d - s - 1),$$

which implies $d \geq t - 0.5 + 0.5w$. On the other hand, $E(A) \geq d - 1 + w$, so $d \leq n + 1 - t - w$.

By $n + 1 - t - w \geq d \geq t - 0.5 + 0.5w$, we know that $0 \leq w \leq 4$. By

$$n - t \geq E(A) \geq d - 1 + w - 1 + \max\{x_{i+1} - x_i, x_1\} - 1, \quad d \geq t - 0.5 + 0.5w,$$

we know that $D \triangleq \max\{x_{i+1} - x_i, x_1\} \le 8.5 - 1.5w$.

Then at most w number of $y_i - y_{i-1}$ is not 1. Take $\lceil \frac{q-w}{2} \rceil$ disjoint pairs (y_i, y_{i+1}) in A_2 such that $y_{i+1} - y_i = 1$. Note that $x_i - x_{i-1}, x_1 \le D$, so if $\lceil \frac{q-w}{2} \rceil \ge D - 2$, then based on selected 1-pairs, we can absorb $x_1, x_2, \cdots, x_p$ one by one. If the sum of absorbed elements

$$\sum_{i=1}^{p} x_i + \lceil \frac{q-w}{2} \rceil \ge \frac{p(p+1)}{2} + \frac{q-w}{2},$$

is not less than $d-2$, then it can absorb all elements in $\bar{A}_2 \cup \{d\}$, implying A reaches our goal , a contradiction. Hence, one of $p^2 + p + q - w \le 2d - 6, q \le 2D + w - 6$ must hold.

If $x_1 + x_2 \ge d + s + 1$, then $x_2 > \frac{d+1+s}{2}$, and consequently

$$y_q > a_2 + d + (t-3) \ge n \quad (\text{for } n \ge 25),$$

a contradiction. Hence $x_1 + x_2 \le d - s - 1$, and by the same reasoning, $x_p + x_{p-1} \le d - s - 1$, which gives $p \le \frac{n+1-s-t-w}{2}, p \le \frac{n+3-D-s-t-w}{2}$.

If $q \le 2D + w - 6$, then $p \ge t - 2D - w + 4$, so

$$n - 1 - s - t - w \ge 2t - 4D - 2w + 8,$$

and

$$n + 3 - D - s - t - w \ge 2t - 4D - 2w + 8,$$

this cannot be possible when $n \ge 32, w \ge 3$ or $D \le 6$; if else, $p \le \sqrt{2(n - 1 - 1.5t)}$.

Similarly, $q - 1 \le y_q - y_1 \le d - s - 1 \le n + 2 - t - s - w - D$.

Combining all above argument, we have either

$$t - 2 = p + q \le \frac{3}{2}(n + 3 - s - t - w - D), \quad 1 \le w \le 2, D \ge 7$$

or

$$t - 2 = p + q \le \sqrt{2(n - 1 - 1.5t)} + n + 1 - t - s - w, \quad w \ge 3,$$

Both of which contradict the assumption $n \ge 22$. $\square$

Finally, we show that this case is not possible. By Claim 2, we choose x, y in A_1 or A_2 such that $z = |d - |x \pm y|| \le s$. Among all admissible pairs, we give priority to those involving a subtraction (i.e., with a minus sign) and, in particular, to the pair for which z is the largest.

Consider the set $B = A - \{a_{j-1}, a_j, x, y\}$, then $|B| \ge t - 4, |\bar{B}| \ge s - 2$ and no element in $\bar{B}$ are in the form of $a_u - a_v$ where $u \ge j + 1, v \le j - 2$ because of our choice of x, y. If $\mathcal{L}(\bar{B}, |\bar{B}| + 2 - m) \ge m + 1$, then

$$n - t + 2 \ge E(A - \{x, y\}) \ge d + (|\bar{B}| + 1 - m)(m + 1),$$

which only holds when $m = 0, 1$ or $|\bar{B}| - 1$ under assumption $n \ge 36$. Let $B^{\uparrow} = \{c_i\}_{i=1}^{h}, c_0 = 0$, $\bar{B}^{\uparrow} = \{b_i\}_{i=1}^{l}$. Then either $b_1 \ge 3$ or $b_{l-1} \ge l + 1$ or $b_l \ge l + 2$, and $X = \{b_1, \cdots, b_{l-2}\}$ is 1-strongly expressive. The situation can be divided into 3 cases as follows:

1. $b_1 \le 2, b_l \ge l+2, b_{l-1} \le l$. Then $\{b_1, \cdots, b_{l-1}\}$ is 1-strongly expressive. By our choice of x, y, for all b_i we have either $|b_i - d| \ge s+1$ or $|b_i - d| \le z$. Moreover, $|b_l - z| \ge l+2 \ge s$ otherwise $X \cup \{|b_l - z|\}$ is 1-strongly expressive, a contradiction. Let $\Delta = \sum_{i=1}^{l-1}(b_i - 1) + (b_l - l - 2) + d - (s+1) + s(\tilde{B} - \{a_{j+1} - a_{j-2}\}) - (|\tilde{B}| - 1)$, then

$$n - t + 2 \ge E(A - \{x, y\}) \ge l + 1 + s + 1 + \Delta, \tag{6}$$

which implies $\Delta \le 7$.

Since $z \le s, |b_l - z| \ge s, b_l \ge 1$, we must have $b_l \ge z + l + 2$, so $z \le \Delta \le 7$. If there is $1 \le i \le l-1$ such that $|b_i - d| \le z$, then $b_i - 1 \ge s - z, \Delta \ge z + s - z > 7$ ($n \ge 34$), a contradiction. Hence, for all $1 \le i \le l-1$, $b_i \le d - s - 1$. If $|b_l - d| \ge s+1$, then $b_l - l - 2 \ge d + s - l - 1 > \Delta$ ($n \ge 30$), a contradiction. So $b_l \le d + z$.

Since $l \ge 7$ ($n \ge 38$), there exist 4 consecutive elements $c_i \sim c_{i+3}$ such that $b_u = c_{i+1} - c_i, b_v = c_{i+3} - c_{i+2}$, $u, v < l$. If $c_{i+2} - c_i \ge d - z$, then $\Delta \ge d - z - 2 + z > 8$, a contradiction. So $2 \le c_{i+2} - c_i \le d - s - 1$, $d \ge s+3$. If $c_{i+3} - c_i \ge d - z$, then $\Delta \ge d - z - 3 + z + d - s - 1 > 8$, a contradiction. So $3 \le c_{i+3} - c_i \le d - s - 1$, $d \ge s + 4$.

If $b_k \ge 4$ for some $k < l$, then $b_l \ge 3 + \sum_{i=1}^{l-1} b_i \ge l + 5$, we have $\Delta \ge 3 + 3 + 3 > 7$, a contradiction. Similarly, there is at most one $b_k (k < l)$ such that $b_k = 3$.

Reconsider $c_i \sim c_{i+3}$, first we absorb all other $b_i (i < l)$ except b_u, b_v (it is admissible since all such elements are no more than 3 and only one possible exception) , the sum reaches at least $l - 3 \ge 4$. Then we absorb $c_{i+2} - c_{i+1} \le 5$ and next $c_{i+3} - c_i$ such that $(c_{i+3} - c_i) - (c_{i+2} - c_{i+1}) \le 6$. It is easy to check that we can absorb in this order. Then the sum reaches at least $l - 3 + 4 = l + 1$, since A does not meet our requirement, we must have $b_l \ge l + 4$. Then we have $\sum_{i=1}^{l-1}(b_i - 1) + s(\tilde{B} - \{a_{j+1} - a_{j-2}\}) - (|\tilde{B}| - 1) \le \Delta - 2 - 3 \le 2$.

Since $l \ge 7$, except $c_i \sim c_{i+3}$, there exist also $c_j \sim c_{j+3}$ such that $j < i$ or $j > i + 3$, and $c_{j+1} - c_j = b_w, c_{j+3} - c_{j+2} = b_x$ with $w, x < l$. Then we can first absorb all other b_k ($k < l$) except b_u, b_v, b_w, b_x (it is admissible since the minimal is no more than 2 and the maximal is no more than 3). The sum reaches at least $l - 5 \ge 2$. Then we absorb $c_{i+2} - c_i, c_{i+3} - c_{i+1}, c_{j+2} - c_j, c_{j+3} - c_{j+1}$ (it is admissible since they are no more than 4). The sum now reaches at least $l + 3$, so $b_l - z \ge l + 6$.

In this case, to let the equality (6) hold, we must have

$$\Delta = 7, \quad z = 0, \quad t \equiv 0 \pmod 2,$$

and no other term except d, b_l contributes to the residue energy of $A - \{x, y\}$. Hence, denote $b_l = c_g - c_{g-1}$,and possible $r \ge 1$ such that $c_{r-1} < a_{j-1}, c_r > a_j$, then $c_k - c_{k-1} = 1$ as long as $k \ne g, r$. Moreover, for all $u, v \in A_1$ or A_2 such that $|u - v| < d$, we must have $|u - v| \le 3$. Hence, for all $1 \le i \le h$, $[i, i+3] \cap \{g, r\} \ne \varnothing$. But when $n \ge 38$, we have $t \ge 18, h \ge 14$, considering $[1, 4], [5, 8], [9, 12]$ yields a contradiction.

2. $b_1 \le 2, b_{l-1} \ge l+1$. Then $b_l - b_{l-1} \le l$, otherwise

$$n - t + 2 \ge E(A - \{x, y\}) \ge 3l + 1 + d,$$

a contradiction when $n \ge 31$. Therefore, $X \cup \{b_l - b_{l-1}\}$ is 1-strongly expressive. If $z \le b_1 + \cdots + b_{l-2} + b_l - b_{l-1} + 2$, then A satisfies our requirement, a contradiction. This means $s \ge z \ge l + 1 + (b_l - b_{l-1}) \ge s - 1$.

If $b_l = b_{l-1}$, then $b_l + b_{l-1} - z \le l$ (which means $X \cup \{b_l + b_{l-1} - z\}$ is 1-strongly expressive, a contradiction) unless $z = b_l = b_{l-1} = l+1, X = I_{l-2}$ or $b_{l-1} \ge l+2$, but the latter will lead to

$$n - t + 2 \ge E(A - \{x, y\}) \ge 2l + 2 + d,$$

a contradiction when $n \ge 38$.

If $b_l - b_{l-1} = 1$, then $z = s = l + 2, X = I_{l-2}$, and for the same reason we have $b_l = l + 2, b_{l-1} = l + 1$.

Hence, we must have $X = I_{l-2}, \max\{z, b_l, b_{l-1}\} \le l + 2$. Among all b_i in X, two of them must be in the form of $b_u = c_i - c_{i-1}, b_v = c_{i+2} - c_{i+1}, c_{i+1} - c_i = 1$; otherwise

$$n - t + 2 \ge E(A - \{x, y\}) \ge 2l + d + l - 4,$$

a contradiction when $n \ge 38$. Let $b'_u = c_{i+1} - c_{i-1}, b'_v = c_{i+2} - c_i$ and replace b_u, b_v. Then the updated X' is 1-strongly expressive with $s(X') = l$, so X' can absorb z, b_l, b_{l-1}, a contradiction.

3. $b_1 \ge 3$. Let $\Delta = \sum_{i=1}^{l}(b_i - 3) + d - (s+1) + s(\tilde{B} - \{a_{j+1} - a_{j-2}\}) - (|\tilde{B}| - 1)$, then

$$n - t + 2 \ge E(A - \{x, y\}) \ge 2l + s + 1 + \Delta,$$

implying $\Delta \le 2$ when $n \ge 34$. There are at least $l - 2 - \Delta > 0$ ($n \ge 30$) elements in $\tilde{B} - \{a_{j+1} - a_{j-2}\}$ are unit. Take one of them as the basis X_0, then it can absorb elements in $\bar{B}$ such that they value no more than 3 and are not broken by our action on choosing unit. Then it at least absorbs $l - 2 - \Delta > 1$ ($n \ge 34$) elements. $s(X_0)$ reaches at least 7. All other elements in $\bar{B}$ are no more than 5, and the broken one's gap is no more than 9, so X_0 can absorb all elements in $\bar{B}$ and a broken gap. Now we have

$$s(X_0) \ge 1 + 3(l-2) + 7 \ge 3s - 4 \ge \max\{d, z\},$$

so A meets our requirement, a contradiction.

5.2 Case 2: $\mathcal{L}(\bar{A}, 3) \ge |\bar{A}|$.

Here, $a_{i+1} - a_i \ge 3$ for all $i \equiv t \pmod 2$. Let $\Delta = s(\bar{A}) - 3|\bar{A}| + s(\tilde{A}) - |\tilde{A}|$. Then

$$n - t \ge E(A) \ge \Delta + 2s,$$

so $\Delta \leq 5$. This means at most 2 elements in $\tilde{A}$ are greater than 2. Suppose that there are h elements in $\tilde{A}$ greater than 2. Then there exists $i \leq 2h+2$ such that $a_i - a_{i-1} \leq 2$. Choose such minimal i. Then for all $1 \leq j \leq i-1$, we have $a_j - a_{j-1} \geq 3$. They take at least $2\lfloor \frac{i-1}{2} \rfloor$ in Δ. Then $\{a_i - a_{i-1}\}$ can absorb elements in $\bar{A}$ which value 3 and generated by numbers in $(a_i, a_t]$. There are at least $\lfloor \frac{t-i}{2} \rfloor - x_1$ such elements, where x_1 is the number of exceptions. After that, the sum will come to at least $a_i - a_{i-1} + 3(\lfloor \frac{t-i}{2} \rfloor - x_1)$. For all possible i and $x_1 \leq 5 - 2\lfloor \frac{i-1}{2} \rfloor$, check that when $n \geq 37$, we have

$$3(\lfloor \frac{t-i}{2} \rfloor - x_1) \geq 4 + (5 - 2\lfloor \frac{i-1}{2} \rfloor - x_1).$$

Then it can absorb $a_{i+1} - a_{i-2}, a_{i-3} - a_{i-4} \cdots$ or $a_{i-2} - a_{i-3}, \cdots$ since all these remaining differences are no more than $a_i - a_{i-1} + 11 - 2\lfloor \frac{i-1}{2} \rfloor - x_1$. Our operation implies that A satisfies our requirement, a contradiction.

6 Proof of Theorem 2

If $1 \in A$ or there exist $x, y \in A$ such that $|x - y| \leq 1$, then removing $\{x, y\}$ (or $\{1\}$) leaves a set whose size satisfies

$$|A \setminus \{x, y\}|,\ |A \setminus \{1\}| \geq c_n - 1 \geq \frac{n-5}{2}.$$

By Theorem 3 we finish the proof.

It remains to consider the case where all consecutive elements of A are at least distance 2 apart; that is, $a_i - a_{i-1} \geq 2$ for all i. In this situation, $1 \notin A$, such a configuration is impossible when $n = 8k+1, 8k+2, \ldots, 8k+5$. Also, we have checked that when $n = 25$, all such sets (see in Table 2) has a signed sum less than 2. Thus, for $n \geq 26$, we only need to analyze the cases $n = 8k+6, 8k+7, 8k+8, 8k+9$ with $k \geq 3$, under the assumption (by induction) that $n \in A$ and $|A| = \lfloor \frac{n}{2} \rfloor$ or $\lfloor \frac{n+1}{2} \rfloor$.

- **Case $n = 8k+6$:** Then $|A| = 4k+3$, and since $A \subseteq \{2, 3, \ldots, 8k+6\}$ with gaps of at least 2, the only possible choice is the even numbers:

  $$A = \{2, 4, 6, \ldots, 8k+6\},$$

  which has a signed sum equaling to 0.

- **Case $n = 8k+7$:** The only admissible sets have the form

  $$A = \{2, 4, \ldots, 2m, 2m+3, 2m+5, \ldots, 8k+7\}$$

  for some m.

  - If $m \geq 8$, we can perform the following reductions:

    $$(2m, 2m-2) \to 2, \quad (2m-2, 2m-4) \to 2, \quad (2, 2m, 2m+3) \to 1.$$

    This yields a new multiset containing 1 along with smaller even numbers and the tail starting from $2m+5$. It can be checked that starting from $\varnothing$, we can absorb all elements in this multiset in non-decreasing order.

- If $2 \leq m \leq 7$, in addition to $(2, 2m, 2m+3) \to 1$,we combine large elements:

$$(8k+7, 8k+5) \to 2, \quad (8k+3, 8k-3) \to 6, \quad (8k+1, 8k-1) \to 2,$$

eventually reducing to a set of the form

$$\{1, 2, 2, 6 \ldots, 4, 2m-2, 2m+5, \ldots, 8k-5\},$$

It can be checked that starting from $\varnothing$, we can absorb all elements in this multiset in non-decreasing order.

- If $m = 0$ or 1, we take pairs for all consecutive odd numbers which are greater than 7. The multiset becomes $\{2, 5, 7, 2 \cdots, 2\}$ or $\{3, 5, 7, 2 \cdots, 2\}$, which has signed sum 0 or 1.

- **Case $n = 8k + 8$:** Here $|A| = 4k + 4$, and the only feasible set with minimal spacing is

$$A = \{2, 4, 6, \ldots, 8k+8\},$$

which has a signed sum equaling to 0.

- **Case $n = 8k + 9$:** Then $|A| = 4k + 4$, and the only possible structure is

$$A = \{2, 4, \ldots, 2m, 2m+3, \ldots, 8k+9\}.$$

The construction proceeds analogously to the $n = 8k + 7$ case, using similar reduction steps to eventually produce a 0-strongly expressive multiset.

Combining all these cases completes the proof.

7 Proof of Theorem 1 Aided by Computer

Suppose that $|A| > c_n$, when $18 \leq n \leq 37$, by the form of c_n, we only need to consider the following cases:

n	19	23	25	27	31	33	35
$\|A\|$	10	11	12	14	15	16	18
	11	12	13	15	16	17	19

Table 1: Possible configurations

Based on Proposition 2 and the discussion in our proof, we use the Algorithm 4 to identify sets that are potential counterexamples. The output is false means the input set A is a possible counterexample. The algorithm combines four complementary screening strategies: (1) direct verification via zero-strong expressiveness, as stated in Algorithm 1 with $r = 0$; (2) relaxed verification after removing critical elements (e.g., 1 or adjacent pairs), as stated in Algorithm 1 with $r = 1$; (3) a structural reduction heuristic for sets that have large gaps, as stated in Algorithm 2; (4) a greedy absorption-and-difference procedure, as stated in Algorithm 3. This layered approach ensures both computational efficiency and a high filter rate.

7.1 Algorithm Description

Algorithm 1 IsStronglyExpressive (X_1, r)

Require: Multiset X_1, parameter $r \in \mathbb{N}$

$s \leftarrow 0$
for each $x \in X_1^{\uparrow}$ **do**
    if $x > s + 1 + r$ **then**
        return false
    end if
    $s \leftarrow s + x$
end for
return true

Algorithm 2 SecondaryScreen (X_2)

Require: Sorted set $X_2^{\uparrow} = [a_1, a_2, \ldots, a_t]$ with $t \geq 4$, denote $a_0 = 0$

$d_{\max} \leftarrow \max(\bar{X}_2)$
Locate index j corresponding to the element(s) in X_2 that generate $d_{\max}$ in $\bar{X}_2$
if no such j exists **then**
    return false
end if
Partition $X_2 \backslash \{a_j, a_{j-1}\}$ into $A_1 = X_2 \cap (0, a_{j-1}), A_2 = X_2 \cap (a_j, n]$.
Initialize candidate list $\mathcal{C} \leftarrow \emptyset$
for all pairs $(x, y) \in (A_1 \times A_1) \cup (A_2 \times A_2)$ with $x > y$ **do**
    Add candidate $(z = x + y,\ x,\ y,\ \delta = |d_{\max} - z|)$ to $\mathcal{C}$
    Add candidate $(z = x - y,\ x,\ y,\ \delta = |d_{\max} - z|)$ to $\mathcal{C}$
end for
if $\mathcal{C} = \emptyset$ **then**
    return false
end if
Sort $\mathcal{C}$ by δ in ascending order
for the first $K = \min(5, |\mathcal{C}|)$ candidates in $\mathcal{C}$ **do**
    Let (z, x, y, δ) be the current candidate
    Construct $Y \leftarrow X_2 \setminus \{a_{j-1}, a_j, x, y\}$
    If $\delta > 0$, append δ to Y
    return IsStronglyExpressive $(Y, 0)$ if the output is **true**
end for
return false

Algorithm 3 GreedyScreen (X_3)

Require: Set $X_3 \subset \mathbb{N}$

Initialize set $\mathcal{R} \leftarrow X_3$, sum $s \leftarrow 0$
while $\mathcal{R} \neq \emptyset$ **do**
    Find largest $a \in \mathcal{R}$ such that $a \leq s + 1$
    if such a exists **then**
        $s \leftarrow s + a$, $\mathcal{R} \leftarrow \mathcal{R} \setminus \{a\}$
        continue
    end if
    Find pair $(x, y) \in \mathcal{R} \times \mathcal{R}$ $(x > y)$ maximizing $d = x - y$ subject to $d \leq s + 1$
    if no such pair exists **then**
        return false
    end if
    $s \leftarrow s + d$, $\mathcal{R} \leftarrow \mathcal{R} \setminus \{x, y\}$
end while
return true

Algorithm 4 IsLikelyGood

Require: Set $A \subset \mathbb{N}$

return IsStronglyExpressive ($\bar{A}, 0$) if the output is **true**.
$A' \leftarrow A$, removed $\leftarrow$ **false**
if $1 \in A'$ **then**
    Remove 1 from A', removed $\leftarrow$ **true**
else if exists adjacent pair (a_i, a_{i+1}) with $a_{i+1} - a_i = 1$ **then**
    Remove both elements from A', removed $\leftarrow$ **true**
end if
if removed **then**
    return IsStronglyExpressive ($\bar{A}', 1$) if the output is **true**
end if
return SecondaryScreen (A) if the output is **true**
return GreedyScreen (A) if the output is **true**
return false

7.2 Remaining Counterexamples After Filtering

After algorithmic elimination for all possible configuration in Table 1, the final potential counterexamples are very few. We finish the discussion manually for each case.

As the above tables suggest, when $n \geq 19$, such examples have no 1-gap. We have checked by hand that all these examples are not true counterexamples.

When $n \leq 18$, we confirm that there is no counterexample by filtering.

Now we give explicit counterexamples when $|A| = c_n$. For each n, it is easy to check that the following $|A_n| = c_n$ are counterexamples:

$$A_1 = \varnothing, A_2 = \{2\}, A_3 = A_4 = \{1, 3\}, A_5 = \{1, 2, 5\};$$

2	4	6	8	10	12	14	16	18	20	22
2	4	6	8	10	12	14	16	18	21	23
2	4	6	8	10	12	14	17	19	21	23
2	4	6	8	10	13	15	17	19	21	23
2	4	6	9	11	13	15	17	19	21	23
2	5	7	9	11	13	15	17	19	21	23

Table 2: Possible counterexamples for $n = 19, 23$.

2	4	6	8	10	12	14	16	18	20	22	24
2	4	6	8	10	12	14	16	18	20	23	25
2	4	6	8	10	12	14	16	19	21	23	25
2	4	6	8	10	12	15	17	19	21	23	25
2	4	6	8	11	13	15	17	19	21	23	25
2	4	7	9	11	13	15	17	19	21	23	25
2	5	7	9	11	13	15	17	19	21	23	25

Table 3: Possible counterexamples for $n = 25, 27$.

2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
2	4	6	8	10	12	14	16	18	20	22	24	26	29	31
2	4	6	8	10	12	14	16	18	20	22	25	27	29	31
2	4	6	8	10	12	14	16	18	21	23	25	27	29	31
2	4	6	8	10	12	14	17	19	21	23	25	27	29	31
2	4	6	8	10	13	15	17	19	21	23	25	27	29	31
2	4	6	9	11	13	15	17	19	21	23	25	27	29	31
2	5	7	9	11	13	15	17	19	21	23	25	27	29	31

Table 4: Possible counterexamples for $n = 31$.

2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32
2	4	6	8	10	12	14	16	18	20	22	24	26	28	31	33
2	4	6	8	10	12	14	16	18	20	22	24	27	29	31	33
2	4	6	8	10	12	14	16	18	20	23	25	27	29	31	33
2	4	6	8	10	12	14	16	19	21	23	25	27	29	31	33
2	4	6	8	10	12	15	17	19	21	23	25	27	29	31	33
2	4	6	8	11	13	15	17	19	21	23	25	27	29	31	33
2	4	7	9	11	13	15	17	19	21	23	25	27	29	31	33
2	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33

Table 5: Possible counterexamples for $n = 33, 35$.

$$A_6 = A_7 = \{1, 4, 5, 6\}, A_8 = A_9 = A_{10} = \{1, 2, 6, 7, 8\}, A_{11} = A_{12} = \{1, 7, 8, 9, 10, 11\};$$

$$A_{13} = A_{14} = A_{15} = \{1, 2, 9, 10, 11, 12, 13\};$$

$$A_{16} = A_{17} = \{1, 2, 3, 12, 13, 14, 15, 16\};$$

$$A_n = \{2, 4, \cdots, 2m\}, m \text{ is the largest number such that } 2m \leq n, 4 \nmid m(m+1). \quad n \geq 18$$

Hence, we have finally substantiated our main results and put a perfect full stop to this problem.