An Improved Pseudopolynomial Time Algorithm for Subset Sum

Jiayi Lian
(joint work with Lin Chen, Yuchen Mao, and Guochuan Zhang)

Zhejiang University

September 19, 2025

Problems

- Subset Sum
  - Given a multi-set X of n positive integers and a target t,
  - asks for a subset $Y \subseteq X$ with maximum $\Sigma(Y)$ that does not exceed t.

$$\Sigma(Y) = \sum_{y \in Y} y$$

- NP-hard.

Current Results

Table 1: Pseudopolynomial Time Algorithm for Subset Sum

Running Time	Reference
$O(nt)$	[Bellman '57]
$O(nt/\log t)$	[Pisinger '03]
$\widetilde{O}(\sqrt{n}t)$	[Koiliaris & Xu '17]
$\widetilde{O}(n+t)$	[Bringmann '17][Jin & Wu '19]
$O(nw)$	[Pisinger '99]
$\widetilde{O}(n+w^{5/3})$	[Polak, Rohwedder& Węgrzycki '21]
$\widetilde{O}(n+w^{3/2})$	[Chen, Lian, Mao & Zhang '24]

Conditional Lower Bound: $t^{1-\epsilon} \cdot n^{O(1)}$ [Abboud et al. '22]

Current Results

Table 1: Pseudopolynomial Time Algorithm for Subset Sum

Running Time	Reference
$O(nt)$	[Bellman '57]
$O(nt/\log t)$	[Pisinger '03]
$\widetilde{O}(\sqrt{n}t)$	[Koiliaris & Xu '17]
$\widetilde{O}(n+t)$	[Bringmann '17][Jin & Wu '19]
$O(nw)$	[Pisinger '99]
$\widetilde{O}(n+w^{5/3})$	[Polak, Rohwedder& Węgrzycki '21]
$\widetilde{O}(n+w^{3/2})$	[Chen, Lian, Mao & Zhang '24]
$\widetilde{O}(n+\sqrt{wt})$	this talk
Conditional Lower Bound: $t^{1-\epsilon} \cdot n^{O(1)}$ [Abboud et al. '22]	

Basic Idea

- Compute $\mathcal{S}(X) := \{\Sigma(Y) : Y \subseteq X\}$

$$\boxed{\Sigma(Y) = \Sigma_{y \in Y} y}$$

- Optimaze: find $s \in \mathcal{S}(X)$ such that s is maximum but $s \leqslant t$.

Basic Idea

- Compute $\mathcal{S}(X) := \{\Sigma(Y) : Y \subseteq X\}$ $\boxed{\Sigma(Y) = \Sigma_{y \in Y} y}$
- Optimaze: find $s \in \mathcal{S}(X)$ such that s is maximum but $s \leqslant t$.

Sumset: $A + B = \{a + b : a \in A \cup \{0\}, b \in B \cup \{0\}\}$.

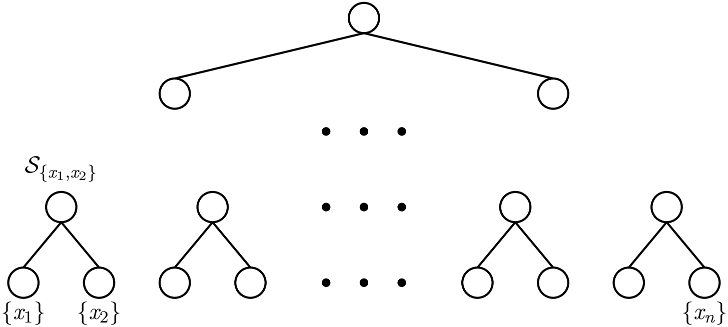
Let $X = \{x_1, ..., x_n\}$, $\mathcal{S}(X) = \{x_1\} + ... + \{x_n\}$.

Basic Idea

- Compute $\mathcal{S}(X) := \{\Sigma(Y) : Y \subseteq X\}$ $\boxed{\Sigma(Y) = \Sigma_{y \in Y} y}$
- Optimaze: find $s \in \mathcal{S}(X)$ such that s is maximum but $s \leqslant t$.

Sumset: $A + B = \{a + b : a \in A \cup \{0\}, b \in B \cup \{0\}\}$.

Let $X = \{x_1, \ldots, x_n\}$, $\mathcal{S}(X) = \{x_1\} + \ldots + \{x_n\}$.

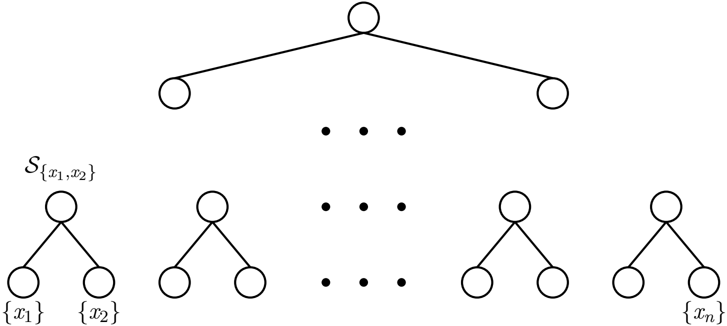


Basic Idea

$\mathcal{S}(X) = \{x_1\} + ... + \{x_n\}$.

Sparse convolution: compute $A + B$ in $\widetilde{O}(|A+B|)$ time.

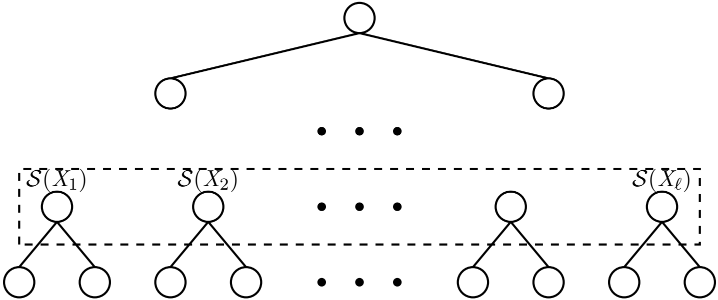
$|\mathcal{S}(X)|$ can be $O(\Sigma x_i) = O(nw)$ or $O(t)$. ($w = \max x_i$)



Basic Idea

Compute $\mathcal{S}(X) := \{\Sigma(Y) : Y \subseteq X\}$. For each layer:

- Sparse Case: $\Sigma|\mathcal{S}(X_i)|$ is small, compute by sparse convolution;
- Dense Case: t is in $\mathcal{S}(X)$ by additive combinatoric. [Bringmann & Wellnitz '21] [Galil & Margalit '91]



Dense Case

Additive combinatorics results from [Szemerédi & Vu '05]

Corollary

There exists a constant c such that: Let $A_1, ..., A_\ell$ be subsets of $[1, u]$. If $\Sigma_{i=1}^{\ell}|A_i| \geqslant cu \log u$, then $A_1 + \cdots + A_\ell$ contains an arithmetic progression of length at least u.

- Common difference $\Delta \leqslant \dfrac{\Sigma_{i=1}^{\ell} \max(A_i)}{u}$.

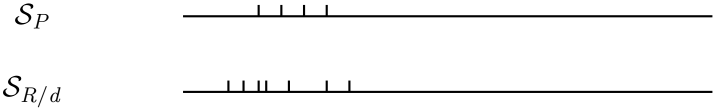
Dense Case

Partition X to G, R, D such that

- $|G| \leqslant k, |R| \leqslant k$,
- R, D divisible by some integer d.
- For any $b \in [1, k]$, $\mathcal{S}_{R/d} \mod b = [1, b-1]$.

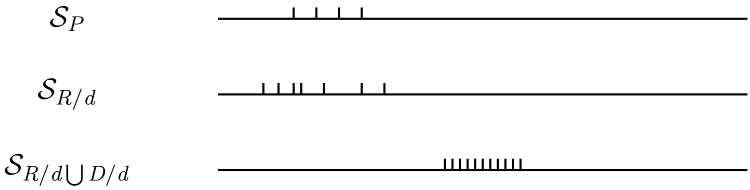
Dense Case

If subset $P \subseteq D/d$ such that
$\mathcal{S}_P$ has an AP with common difference $\Delta \leqslant k$,



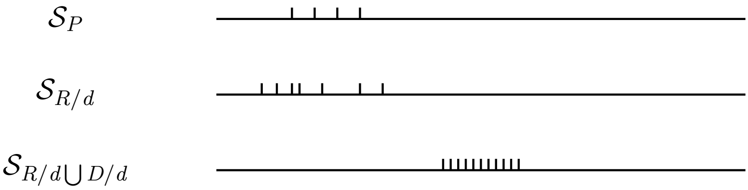
Dense Case

If subset $P \subseteq D/d$ such that
$\mathcal{S}_P$ has an AP with common difference $\Delta \leqslant k$,



Dense Case

If subset $P \subseteq D/d$ such that
$\mathcal{S}_P$ has an AP with common difference $\Delta \leqslant k$,



We have

$$[\sigma(P) + \sigma(R), \sigma(D) - \sigma(P) - \sigma(R)] \subseteq \mathcal{S}_{R/d \bigcup D/d}.$$

If $\sigma(P) + \sigma(R) \leqslant (t - \Sigma(G))/d$,

$$t \in \mathcal{S}(X) \iff t \in d \cdot [\sigma(P) + \sigma(R), \sigma(D) - \sigma(P) - \sigma(R)] + \mathcal{S}_G.$$

Dense Case

To make $\Delta = \dfrac{\Sigma_{i=1}^{\ell} \max(\mathcal{S}(X_i))}{u} \leqslant k$ (also $\sigma(P) + \sigma(R) \leqslant t - \sigma(G)$),

Dense Case

To make $\Delta = \dfrac{\Sigma_{i=1}^{\ell} \max(\mathcal{S}(X_i))}{u} \leqslant k$ (also $\sigma(P) + \sigma(R) \leqslant t - \sigma(G)$),

- By Color Coding, $\Sigma_{i=1}^{\ell} \max(\mathcal{S}(X_i)) = \widetilde{\Theta}(t)$

Dense Case

To make $\Delta = \frac{\Sigma_{i=1}^{\ell} \max(\mathcal{S}(X_i))}{u} \leqslant k$ (also $\sigma(P) + \sigma(R) \leqslant t - \sigma(G)$),

- By Color Coding, $\Sigma_{i=1}^{\ell} \max(\mathcal{S}(X_i)) = \widetilde{\Theta}(t)$
- Running time
  - Compute $\mathcal{S}_G$ and $\mathcal{S}_R$: $\widetilde{O}(|G| \cdot w) = \widetilde{O}(wk)$
  - Compute $\mathcal{S}_{D/d}$ when sparse: $\widetilde{O}(\Sigma_{i=1}^{\ell} |\mathcal{S}(X_i)|) = \widetilde{O}(u)$

Dense Case

To make $\Delta = \frac{\Sigma_{i=1}^{\ell} \max(\mathcal{S}(X_i))}{u} \leqslant k$ (also $\sigma(P) + \sigma(R) \leqslant t - \sigma(G)$),

- By Color Coding, $\Sigma_{i=1}^{\ell} \max(\mathcal{S}(X_i)) = \widetilde{\Theta}(t)$
- Running time
  - Compute $\mathcal{S}_G$ and $\mathcal{S}_R$: $\widetilde{O}(|G| \cdot w) = \widetilde{O}(wk)$
  - Compute $\mathcal{S}_{D/d}$ when sparse: $\widetilde{O}(\Sigma_{i=1}^{\ell} |\mathcal{S}(X_i)|) = \widetilde{O}(u)$
- Minimize: $wk = u = \sqrt{wt}$

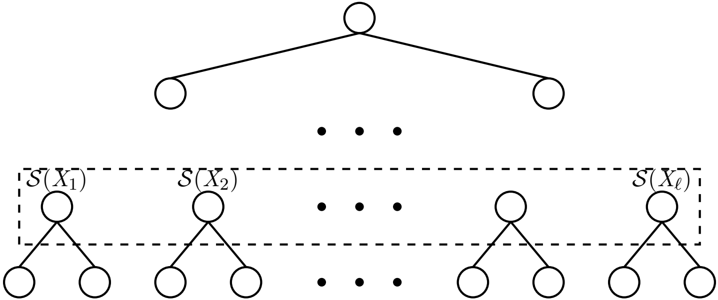
End of Story?

Challenges

Recall $\mathcal{S}(X_1), ..., \mathcal{S}(X_\ell)$ should be subsets of $[1, u]$.
Then if $\Sigma_{i=1}^{\ell}|\mathcal{S}(X_i)| \geqslant cu \log u$, there is an AP.

Challenges

Recall $\mathcal{S}(X_1), ..., \mathcal{S}(X_\ell)$ should be subsets of $[1, u]$.
Then if $\Sigma_{i=1}^{\ell}|\mathcal{S}(X_i)| \geqslant cu\log u$, there is an AP.



In high layer, $\max(\mathcal{S}(X_i)) \geqslant \sqrt{wt}$.

Probability Bounds

From Bernstein's inequality:

Corollary

Let A be a multi-set of k non-negative integers. Let $A^* \subseteq A$. Let B be a multi-set of s integers randomly sampled from A without replacement. For any $c \geqslant 1$,

$$\mathbf{Pr}\left(\left|\Sigma(B \cap A^*) - \frac{s}{k}\Sigma(A^*)\right| > 4c\sqrt{|A^*|}\max(A^*)\right) \leqslant \exp(-c).$$

Probability Bounds

- Suppose $Y \subseteq X$ that $\Sigma(Y) = t$. Let $Y_j = Y \cap [2^j, 2^{j+1}]$.
- Randomly permute X. In layer h, X_i is a set of 2^h integers randomly sampled from X without replacement.
- With high probability,

$$\left|\Sigma(X_i \cap Y_j) - \frac{2^h}{n}\Sigma(Y_j)\right| \leqslant c\sqrt{|Y_j| \max(Y_j)} \leqslant c\sqrt{t \cdot 2^j} \leqslant c\sqrt{wt}$$

- With high probability,

$$\left|\Sigma(X_i \cap Y) - \frac{2^h}{n}\Sigma(Y)\right| \leqslant \sum_{j=1}^{\log w} c\sqrt{|Y_j| \max(Y_j)} \leqslant c \log w \sqrt{wt}$$

In layer h, we can just compute

$$\mathcal{S}(X_j) \bigcap \left[\frac{2^h}{n}t - c\log w\sqrt{wt}, \frac{2^h}{n}t + c\log w\sqrt{wt}\right].$$

Corollary

There exists a constant c such that: Let $A_1, ..., A_\ell$ be subsets of $[C, C+u]$. If $\Sigma_{i=1}^{\ell}|A_i| \geqslant cu\log u$, then $A_1 + \cdots + A_\ell$ contains an arithmetic progression of length at least u.

Open Questions

- Does Subset Sum (or Partition) have an exact algorithm in time $\widetilde{O}(n + w)$?
- Does Partition admit a deterministic approximation scheme in time $\widetilde{O}(n + 1/\varepsilon)$?
- Does Subset Sum admit a deterministic weak approximation scheme in time $\widetilde{O}(n + 1/\varepsilon)$?

Thank you!