Approximating Knapsack and Partition via Dense Subset Sums

Jan Klimczak, Szymon Wojtulewicz

Uniwersytet Jagielloński

26.10.2023

Abstract

Knapsack

Knapsack can be $(1-\varepsilon)$-approximated in $\tilde{O}(n+(1/\varepsilon)^{2.2})$ time, improving the previous $\tilde{O}(n+(1/\varepsilon)^{2.25})$ by Jin (ICALP'19). There is a known conditional lower bound of $(n+1/\varepsilon)^{2-o(1)}$ based on $(\min,+)$-convolution hypothesis.

Partition

Partition can be $(1-\varepsilon)$-approximated in $\tilde{O}(n+(1/\varepsilon)^{1.25})$ time, improving the previous $\tilde{O}(n+(1/\varepsilon)^{1.5})$ by Bringmann and Nakos (SODA'21). There is a known conditional lower bound of $(1/\varepsilon)^{1-o(1)}$ based on Strong Exponential Time Hypothesis.

Problem statements

Knapsack

The input is a list of n items $(p_1, w_1), \ldots, (p_n, w_n) \in \mathbb{N} \times \mathbb{N}$ together with a knapsack capacity $W \in \mathbb{N}$, and the optimal value is

$$\mathrm{OPT} := \max_{J \subseteq [n]} \left\{ \sum_{j \in J} p_j \,\middle|\, \sum_{j \in J} w_j \leq W \right\}.$$

Partition

The input is a list of n integers $x_1, \ldots, x_n \in \mathbb{N}$, and the optimal value is

$$\mathrm{OPT} := \max_{J \subseteq [n]} \left\{ \sum_{j \in J} x_j \,\middle|\, \sum_{j \in J} x_j \leq \frac{1}{2} \sum_{i \in [n]} x_i \right\}.$$

Approximation

Approximation

Given a Knapsack (or a Partition) instance and a parameter $\varepsilon \in (0, 1)$, an $(1 - \varepsilon)$-*approximation algorithm* is required to output a number SOL such that $(1 - \varepsilon)\mathrm{OPT} \leq \mathrm{SOL} \leq \mathrm{OPT}$.

In both problems, we can assume $n = O(\varepsilon^{-4})$ and hence $\log n = O(\log \varepsilon^{-1})$. For larger n, Lawler's algorithm for Knapsack in $O(n \log \frac{1}{\varepsilon} + (\frac{1}{\varepsilon})^4)$ time is already near-optimal.

Approximation

Definition

Given a Knapsack (or a Partition) instance and a parameter $\varepsilon \in (0,1)$, an $(1-\varepsilon)$-*approximation algorithm* is required to output a number SOL such that $(1-\varepsilon)\mathrm{OPT} \leq \mathrm{SOL} \leq \mathrm{OPT}$.

In both problems, we can assume $n = O(\varepsilon^{-4})$ and hence $\log n = O(\log \varepsilon^{-1})$. For larger n, Lawler's algorithm for Knapsack in $O(n \log \frac{1}{\varepsilon} + (\frac{1}{\varepsilon})^4)$ time is already near-optimal.

Definitions

Definition 2.2

For functions $\tilde{f}, f$, and real numbers $t, \Delta \in \mathbb{R}_{\geq 0}, \delta \in [0,1)$, we say that $\tilde{f}$ is a $(1-\delta, \Delta)$ approximation of f up to t, if

1. $\tilde{f}(w) \leq f(w)$ holds for all $w \geq 0$
2. $\tilde{f}(w) \geq (1-\delta)f(w) - \Delta$ holds whenever $f(w) \leq t, w \geq 0$

Sumsets and Subset Sums

Definition

In a multiset A, an element a could appear multiple times (the number of times it appears is the *multiplicity* of a in A). We use $A \uplus B$ to denote union without removing duplicates (i.e., possibly resulting in a multiset).

Definition

For a multiset $Y \subset \mathbb{N}$, let $\Sigma(Y) = \sum_{y \in Y} y$ denote the sum of its elements (without removing duplicates).

Definition

For a multiset $X \subset \mathbb{N}$, let $\mathcal{S}(X) = \{\Sigma(Y) : Y \subseteq X\}$ be the set of its subset sums, and let $\mathcal{S}(X; t) = \mathcal{S}(X) \cap [0, t]$ be the set of its subset sums up to t.

Approximation for Integer Sets

Definition

For integer sets $A, B \subseteq \mathbb{N}$, and a real numbers $\delta \in [0, 1)$, we say that A is a $(1-\delta)$*-multiplicative approximation of B*, if

1. for every $b \in B$, there exists $a \in A$ such that $(1-\delta)b \leq a \leq b$, and,
2. for every $a \in A$, there exists $b \in B$ such that $(1-\delta)b \leq a \leq b$.

Definition

For integer sets $A, B \subseteq \mathbb{N}$, and a real number $\Delta \in \mathbb{R}_{\geq 0}$, we say that A is a Δ*-additive approximation of B up to t*, if

1. for every $b \in B$, there exists $a \in A$ such that $b - \Delta \leq a \leq b$, and,
2. for every $a \in A$, there exists $b \in B$ such that $b - \Delta \leq a \leq b$.

Approximation for Integer Sets

Definition

For integer sets $A, B \subseteq \mathbb{N}$, and real numbers $\Delta \in \mathbb{R}_{\geq 0}, \delta \in [0, 1)$, we say that A is a $(1 - \delta, \Delta)$ *approximation of* B, if

1. for every $b \in B$, there exists $a \in A$ such that $(1 - \delta)b - \Delta \leq a \leq b$, and,
2. for every $a \in A$, there exists $b \in B$ such that $(1 - \delta)b - \Delta \leq a \leq b$.

$(1, \Delta)$ approximation is a Δ-*additive approximation*, and $(1 - \delta, 0)$ approximation ia a $(1 - \delta)$-*multiplicative approximation*, or simply $(1 - \delta)$ *approximation*.

Approximation for Integer Sets up to t

Definition

For integer sets $A, B \subseteq \mathbb{N}$, and real numbers $t, \Delta \in \mathbb{R}_{\geq 0}, \delta \in [0, 1)$, we say that A is a $(1 - \delta, \Delta)$ *approximation of B up to t*, if

1. for every $b \in B \cap [0, t]$, there exists $a \in A$ such that $(1 - \delta)b - \Delta \leq a \leq b$, and,
2. for every $a \in A$, there exists $b \in B$ such that $(1 - \delta)b - \Delta \leq a \leq b$.

One can assume $A \subseteq \mathbb{N} \cap [0, t]$ in this case without loss of generality.

Proposition 2.4

For $i \in \{1, 2\}$, suppose A_i is a $(1-\delta, \Delta_i)$ approximation of $\mathcal{S}(X_i)$ up to t. Then, $(A_1 + A_2) \cap [0, t]$ is a $(1-\delta, \Delta_1 + \Delta_2)$ approximation of $\mathcal{S}(X_1 \uplus X_2)$ up to t.

Proposition 2.5

For $i \in \{1, 2\}$, suppose $\tilde{f}_i$ is a $(1-\delta, \Delta_i)$ approximation of the profit function f_{I_i} up to t. Then $(\tilde{f}_1 \oplus \tilde{f}_2)$ is a $(1-\delta, \Delta_1 + \Delta_2)$ approximation of $f_{I_1 \uplus I_2}$ up to t.

Knapsack

Lemma 3.2

Given a list I of items $(p_1, w_1), \dots, (p_n, w_n)$ with weights $w_i \in \mathbb{N}^+$ and profits p_i being multiples of ε in the interval $[1, 2)$, one can $(1 - \varepsilon)$-approximate the profit function f_I up to B in $\tilde{O}(n + \varepsilon^{-2} B^{1/3} / 2^{\Omega(\sqrt{log(1/\varepsilon)})})$

Lemma 3.3

https://www.overleaf.com/project/65352a0d01190f831e160cb6 Given a list I of items $(p_1, w_1), \dots, (p_n, w_n)$ with weights $w_i \in \mathbb{N}^+$ and profits p_i being multiples of ε in the interval $[1, 2)$, if there are only m distinct profit values p_i, then one can $(1 - \varepsilon)$-approximate the profit function f_I in $\tilde{O}(n + \varepsilon^{-3/2} m / 2^{\Omega(\sqrt{log(1/\varepsilon)})})$

Lemma 3.4

Let $f_1, \ldots, f_m$ be monotone step functions with total complexity $O(n)$ and ranges contained in $\{0\} \cup [A, B]$. Then we can compute a monotone step function that has complexity $\tilde{O}(\frac{1}{\varepsilon} log B/A)$ and $(1 - O(\varepsilon))$-approximates $f_1 \oplus \ldots \oplus f_m$ in $O(n) + \tilde{O}((\frac{1}{\varepsilon})^2 m / 2^{\Omega(\sqrt{log(1/\varepsilon)})} log B/A)$ time.

Knapsack

Problem 1

Assume $\varepsilon \in (0, 1/2)$ and $1/\varepsilon \in \mathbb{N}^+$. Given a list I of items $(p_1, w_1), \ldots, (p_n, w_n)$ with weights $w_i \in \mathbb{N}$ and profits p_i being multiples of ε in the interval $[1, 2)$, compute a profit function that $(1-\varepsilon)$-approximates f_I up to $2/\varepsilon$.

Lemma 3.1

If for some $c \geq 2$, Problem 1 can be solved in $\tilde{O}(n + 1/\varepsilon^c)$ time, then $(1-\varepsilon)$-approximating Knapsack can also be done in $\tilde{O}(n + 1/\varepsilon^c)$ time.

Knapsack

Lemma 3.6

Given a list I of $n = O(1/\varepsilon)$ items with p_i being multiples of ε in interval $[1, 2)$, one can compute in $\tilde{O}(n^{4/5}\varepsilon^{-7/5}/2^{\Omega(\sqrt{log(1/\varepsilon)})}$ time a profit function that $(n\varepsilon)$-additively approximates f_I.

Definition

We say a monotone step function is p-uniform if its function values are $0, p, 2p, ..., lp$ for some l. A p-uniform function is said to be *pseudo-concave*, if the differences of consecutive x-breakpoints are nondecreasing from left to right

Definition

For a set Δ of numbers, we say that p is Δ-multiple if it is a multiple of δ for some $\delta \in \Delta$

Knapsack

Lemma 3.10

Let $f_1, ..f_m$ be monotone step functions with ranges contained in $[0, B]$. Let $\Delta \subset [\delta, 8\delta]$. If every f_i is p_i-uniform and pseudo-concave for some $p_i \in [1, 2]$ which is a Δ-multiple, then we can compute a monotone step function that $O(|\Delta|\delta)$-additively approximates $min\{f_1 \oplus ... \oplus f_m, B\}$ in $\tilde{O}(Bm/\delta)$ time.

Lemma 3.11

For parameters $0 < \varepsilon < \delta < 1/2$, let $r = \lceil log_{1+\varepsilon}(1 + 2\delta) \rceil = O(\delta/\varepsilon)$, and define $a_i = \delta(1 + \varepsilon)^i$ for $0 \leq i \leq r + 1$. Let $\Delta = \{a_i\}$ be the set of a_i. Then for any $t \in [1, 2]$, there is a multiple of some a_i in the range $[t, t + 2\varepsilon]$. Thus, every real number in $[1, 2]$ can be approximated by a Δ-multiple with $O(\varepsilon)$ additive error, where $|\Delta| = r + 2 = O(\delta/\varepsilon)$ and all elements in Δ are within $[\delta, 8\delta]$.

Knapsack

Lemma 3.6

Given a list I of $n = O(1/\varepsilon)$ items with p_i being multiples of ε in interval $[1, 2)$, one can compute in $\tilde{O}(n^{4/5}\varepsilon^{-7/5}/2^{\Omega(\sqrt{log(1/\varepsilon)})}$ time a profit function that $(n\varepsilon)$-additively approximates f_I.

Knapsack

Claim 3.12

We can partition elements of I into $\Theta(\Delta_1)$ groups $G_1, ..., G_k$, each of size $O(n/\Delta_1)$, while all elements within group G_i are $(1+\varepsilon)$-approximated by multiples of p_i for some $p_i = \Theta(\Delta_1 \varepsilon)$.

From now on, assume that $G_1, ..., G_k$ are groups satisfying conditions in Claim 3.12.
We now randomly partition $\{1, 2, ..., k\}$ into Δ_0 parts $I_1, ... I_{\Delta_0}$, by assigning each $1 \leq i \leq k$ into some I_j independently and uniformly. Then, set $X_j = \bigcup_{i \in I_j} G_i$.

Claim 3.13

With probability $\geq 3/4$, $|I_j| = O(\Delta_1/\Delta_0)$, and hence $|X_j| \leq O(n/\Delta_0)$

Knapsack

Claim 3.14

We can approximate $\oplus_{x \in X_j} f_x$ with additive error $O(n\varepsilon/\Delta_0)$ for all $1 \leq j \leq \Delta_0$ in $\tilde{O}(n^2\varepsilon^{-1}/(\Delta_0\Delta_1)) = \tilde{O}(n^{4/5}\varepsilon^{-7/5}/2^{\Omega(\sqrt{log(1/\varepsilon)})})$

Claim 3.15

Assume $i \leq log_2\Delta_0$ and $|S_1| = |S_2| = 2^i$, where $S_1, S_2 \subseteq \{1, 2, ...\Delta_0\}$ and $S_1 \bigcap S_2 = \emptyset$. Assume that A_1 is an approximation of $F(S_1)$ with additive error err_1, A_2 is an approximation of $F(S_2)$ with additive error err_2. Then with probability $\geq (1 - 1/(5\Delta_0))$, we can compute an approximation of $F(S_1 \bigcup S_2$ with additive error $err_1 + err_2 + O(2^{0.9i}n\varepsilon/\Delta_0)$ in time $O(\varepsilon^{-2}\Delta_0^{0.5}/(\Delta_1^{0.5}2^{\Omega(\sqrt{log(1/\varepsilon)})})$

Knapsack

Lemma 2.6

Let n distinct positive integers $X = \{x_1, ..., x_n\} \subseteq [l, 2l]$ be given, where $l = o(n^2/log_n)$. Then, for a universal constant $c \geq 1$, for every $cl^2/n \leq t \leq \sum(X)/2$, there exists $t' \in \mathcal{S}(X)$ such that $0 \leq t' - t \leq 8l/n$.

Lemma 3.5

Given a list I of n items with p_i being multiples of ε in interval $[1, 2)$, and integer $1 \leq m \leq n$ with $m = O(1/\varepsilon)$, one can compute in $O(n + \varepsilon^{11/5}/2^{\Omega(\sqrt{log(1/\varepsilon)})})$ time a profit function that $(m\varepsilon)$-additively approximates f_I up to $2m$.

From now on, assume that items are sorted by non-decreasing order of efficiency, $p_1/w_1 \geq p_2/w_2 \geq ... \geq p_n/w_n$.

Definition 3.7

For $1 \leq i \leq n$, let $D(i) = min_J C([i] \setminus J)$, where the minimization is over all subsets $J \subseteq [i]$ with $|J| \leq 2m$, and $C([i] \setminus J)$ denote the number of distinct values in $\{p_j : j \in [i] \setminus J\}$

Observation 3.8

For all $2 \leq i \leq n$, $0 \leq D(i) - D(i-1) \leq 1$.

Lemma 3.9

Let $S \subseteq [n]$ be any item set with total profit $\sum_{s \in S} p_s \leq 2m$. Let $B := 9c\varepsilon^{-1}/\Delta$, where $c \geq 1$ is the universal constant in Lemma 2.6. Then, there exists an item set $\tilde{S} \subseteq [n]$, such that the total profit $\tilde{p}$ contributed by items $[n] \setminus [i]$ in $\tilde{S}$ satisfies:

1. $$\tilde{p} := \sum_{s \in \tilde(S) \cap ([n] \setminus [i])} p_s \leq B$$
2. $$\sum_{s \in \tilde{S}} p_s \geq (1 - \varepsilon) \sum_{s \in S} p_s$$
3. $$\sum_{s \in \tilde{S}} w_s \leq \sum_{s \in S} w_s$$

Knapsack

Lemma 3.9

We define $\tilde{S} \subseteq [n]$ as the maximizer of

$$\sum_{s \in \tilde{S} \cap [i]} p_s + \sum_{s \in \tilde{S} \cap ([n] \setminus [i])} (1 - \varepsilon) p_s$$

Knapsack

Lemma 3.5

Given a list I of n items with p_i being multiples of ε in interval $[1, 2)$, and integer $1 \leq m \leq n$ with $m = O(1/\varepsilon)$, one can compute in $O(n + \varepsilon^{11/5}/2^{\Omega(\sqrt{log(1/\varepsilon)})})$ time a profit function that $(m\varepsilon)$-additively approximates f_I up to $2m$.

Recall that $i \in \{1, 2, ..., n\}$ is the maximum such that $D(i) \leq \Delta$, which can be found with binary search in $\tilde{O}(n)$. Let $J \subset [i]$ with $|J| \leq 2m$ be the minimizer of $D(i)$.

- Use Lemma 3.6 to $(2m\varepsilon)$-additively approximate f_J in $O(m^{4/5}\varepsilon^{-7/5})$
- By definition of i, items in $[i] \setminus J$ have no more that Δ distinct profit values. Use Lemma 3.3 to find $(1 - \varepsilon)$-aaproximation of $f_{[i]\setminus J}$ in $\tilde{O}(\varepsilon^{-17/8}$.
- Use Lemma 3.2 to find $(1 - \varepsilon)$-approximation of $f_{[n]\setminus[i]}$ in $\tilde{O}(B^{1/3}\varepsilon^{-2})$.

Partition

Theorem 1.2

There is a deterministic algorithm for $(1-\varepsilon)$-approximating Partition with running time

$$\tilde{O}\left(n+\varepsilon^{-5/4}\right).$$

Partition

High-level outline

- Reduce the Partition problem to constructing an additive approximation.
- Use *densified* FFT when merging two approximations. The result will have additional additive error.
- Transform an additive approximation algorithm to a multiplicative one.
- Recursively merge the approximations with cumulative multiplicative error.
- Use previous works combined with the new results to construct the additive approximation

Many steps in this work revolve around transforming additive approximations to multipicative ones and vice versa.

Reduction

Problem 2

Assume $\varepsilon \in (0, 1/2)$ and $1/\varepsilon \in \mathbb{N}^+$. Given a set X of n *distinct* integers in the interval $[1/\varepsilon, 2/\varepsilon)$, compute a set $A \subset \mathbb{N}$ that n-additively approximates $\mathcal{S}(X)$.

Lemma 4.1 (Tedious reduction)

If for some $c \geq 1$, Problem 2 can be solved in $\tilde{O}(n + 1/\varepsilon^c)$ time, then $(1-\varepsilon)$-approximating *Partition* can also be solved in $\tilde{O}(n + 1/\varepsilon^c)$ time.

Lemma 4.1

Problem 2

Assume $\varepsilon \in (0, 1/2)$ and $1/\varepsilon \in \mathbb{N}^+$. Given a set X of n *distinct* integers in the interval $[1/\varepsilon, 2/\varepsilon)$, compute a set $A \subset \mathbb{N}$ that n-additively approximates $\mathcal{S}(X)$.

Lemma 4.1 (Tedious reduction)

If for some $c \geq 1$, Problem 2 can be solved in $\tilde{O}(n + 1/\varepsilon^c)$ time, then $(1 - \varepsilon)$-approximating *Partition* can also be solved in $\tilde{O}(n + 1/\varepsilon^c)$ time.

Lemma 4.2

Lemma 4.2

Let $\delta \in (0, 1/2)$, and $\ell, d, t, \Delta \in \mathbb{N}^+$ such that $d \leq \ell \leq t$.
Let $X_1, X_2 \subseteq \mathbb{N}^+ \cap [\ell, \ell + d]$ be two integer sets. Given $A_1, A_2 \subset \mathbb{N}$ as input where for $i \in \{1, 2\}$, A_i is an $(1 - \delta)$ approximation of $\mathcal{S}(X_i)$ up to t, one can compute a set $A \subset \mathbb{N}^+$ of size $|A| \leq Z$ that $(1 - \delta, \Delta - 1)$-approximates $\mathcal{S}(X_1 \uplus X_2)$ up to t, with time complexity:

1. $\tilde{O}(\lceil \frac{t}{\Delta} \rceil + |A_1| + |A_2|)$
2. $\tilde{O}(\frac{t}{\ell} \cdot \lceil \frac{td}{\ell\Delta} \rceil + |A_1| + |A_2|)$
3. $\tilde{O}(Z + |A_1| + |A_2|)$, where $Z \leq O\left(\min\left\{\lceil \frac{t}{\Delta} \rceil,\ \frac{t}{\ell} \cdot \lceil \frac{td}{\ell\Delta} \rceil\right\}\right)$.

3 follows immediately from 1 and 2.

Lemma 4.2

Lemma (1D-FFT)

For a number c and a set X, define $c \cdot X = \{cx : x \in X\}$. For two sets X, Y, define their *sumset* $X + Y = \{x + y : x \in X, y \in Y\}$. Given sets $X \subseteq [n]$, $Y \subseteq [n]$, the sumset $X + Y$ can be computed in $O(n \log n)$ time using FFT.

Proposition 2.4

For $i \in \{1, 2\}$, suppose A_i is a $(1 - \delta, \Delta_i)$ approximation of $\mathcal{S}(X_i)$ up to t. Then, $(A_1 + A_2) \cap [0, t]$ is a $(1 - \delta, \Delta_1 + \Delta_2)$ approximation of $\mathcal{S}(X_1 \uplus X_2)$ up to t.

Lemma 4.2

Proof of Lemma 4.2.1

Let $\bar{\Delta} := \lceil \Delta/2 \rceil$. For $i \in \{1, 2\}$, by rounding the integers in A_i down to multiples of $\bar{\Delta}$, we obtain set $A'_i \subset \bar{\Delta} \cdot \mathbb{N}$ that $(\bar{\Delta} - 1)$-additively approximates A_i. Then, since $A'_i \subseteq [0, t]$, their sumset $A'_1 + A'_2$ can be computed by FFT in $\tilde{O}(\lceil t/\bar{\Delta} \rceil)$ time. Note that $A := A'_1 + A'_2$ is a $(1 - \delta, \Delta - 1)$-approximation of $\mathcal{S}(X_1 \uplus X_2)$ up to t.

Lemma 4.2

Lemma 2.1(2D-FFT)

Given two sets $A_1, A_2 \subseteq [n] \times [m]$, one can compute

$$A_1 + A_2 := \{(x_1 + x_2, y_1 + y_2) : (x_1, y_1) \in A_1, (x_2, y_2) \in A_2\}$$

in $O(nm \log(nm))$ time deterministically.

Lemma 4.2

Proof of Lemma 4.2.2

Every $s \in \mathcal{S}(X_i; t)$ can be expressed as $s = k\ell + b'$, so *"corresponding"* $a \in A_i$ can be expressed as $a = k\ell + b$. Round b down to integer multiples of $\bar{\Delta}$:

$$a' = k\ell + j\bar{\Delta},$$

for some $k \in \mathbb{N} \cap [0, t/\ell]$ and $j \in \mathbb{Z} \cap [-1 - s\delta/\bar{\Delta}, dt/(\ell\bar{\Delta})]$. We have obtained A'_1 which is a $(\bar{\Delta} - 1)$-additive approximation of A_i. Using this 2-dimensional (k, j) representation of A'_i, compute $A'_1 + A'_2$ using 2D FFT.

Observe that $\delta \leq O(d/\ell)$

Lemma 4.3

Let $\delta, \delta_0 \in (0, 1/2)$, and $\ell, d, T \in \mathbb{N}^+$ such that $d \leq \ell \leq T$. Let $X_1, X_2 \subseteq \mathbb{N}^+ \cap [\ell, \ell + d]$ be two integer sets. Given $A_1, A_2 \subset \mathbb{N}$ as input where for $i \in \{1, 2\}$, A_i is an $(1-\delta)$ approximation of $\mathcal{S}(X_i)$ up to T, one can compute a set $A \subset \mathbb{N}^+$ of size $|A| \leq Z$ that $(1-\delta-\delta_0)$-approximates $\mathcal{S}(X_1 \uplus X_2)$ up to T, in $\tilde{O}\big(Z + (|A_1| + |A_2|)\log(2T/\ell)\big)$ time, where

$$Z \leq O\left(\min\left\{\frac{\log(2T/\ell)}{\delta_0},\ \frac{T}{\ell} \cdot \left\lceil \frac{d}{\ell\delta_0} \right\rceil\right\}\right).$$

Lemma 4.3

Algorithm outline for Lemma 4.3

- Initialize set $A = \{0\}$.
- For each r being integer powers of 2 such that $\ell/6 \leq r \leq T$:
  - apply Lemma 4.2 to A_1 and A_2 with $t := 6r$ and $\Delta := \lceil \delta_0 r \rceil$, and obtain a set $A_r \subseteq \mathbb{N} \cap [0, 6r]$
  - Insert all elements in $A_r \cap [r, 6r]$ into A

On every step A_r $(1 - \delta, \lceil \delta_0 r \rceil - 1)$-approximates $\mathcal{S}(X_1 \uplus X_2)$ up to $6r$.

Proof of Lemma 4.3

For every $a \in A_r \cap [r, 6r]$, there exists $s \in \mathcal{S}(X_1 \uplus X_2)$ such that $a \leq s$ and

$$a \geq (1-\delta)s - (\lceil \delta_0 r \rceil - 1) > (1-\delta)s - \delta_0 r \geq (1 - \delta - \delta_0)s$$

For every $s \in \mathcal{S}(X_1 \uplus X_2; T)$, let r be a power of two such that $3r \leq s \leq 6r$. Then there exists $a \in A_r$ such that $a \leq s \leq 6r$ and

$$a \geq (1-\delta)s - (\lceil \delta_0 r \rceil - 1) \geq (1-\delta)s - \delta_0 r \geq s/2 - r/2 \geq r,$$

so $a \in A_r \cap [r, 6r]$ and hence will be included in A, and $a \geq (1 - \delta_0 - \delta)s$.

Corollary 4.4

Corollary 4.4

Let $\delta, \delta_0 \in (0, 1/2)$, and $\ell, d \in \mathbb{N}^+$ such that $d \le \ell$. Let $X_1, X_2 \subseteq \mathbb{N}^+ \cap [\ell, \ell + d]$ be two integer sets of total size $|X_1| + |X_2| = n$. Given $A_1, A_2 \subset \mathbb{N}$ as input where for $i \in \{1, 2\}$, A_i is an $(1 - \delta)$ approximation of $\mathcal{S}(X_i)$, one can compute a set $A \subset \mathbb{N}^+$ of size $|A| \le Z$ that $(1 - \delta_0 - \delta)$-approximates $\mathcal{S}(X_1 \uplus X_2)$, in $\tilde{O}\big(Z + (|A_1| + |A_2|) \log n\big)$ time, where

$$Z \le O\left(\min\left\{\frac{1}{\delta_0},\ \frac{nd}{\ell\delta_0} + n\right\} \cdot \log n\right).$$

Proof

Immediately follows from Lemma 4.3 by setting $T = n \cdot (\ell + d)$, which is an upper bound on the largest element of $\mathcal{S}(X_1 \uplus X_2)$.

Lemma 4.5

Lemma 4.5

Given an integer set $X \subseteq \mathbb{N}^+ \cap [\ell, 2\ell]$ of n integers, one can compute a set $A \subset \mathbb{N}^+$ that $(1-\delta)$-approximates $\mathcal{S}(X)$, in $\tilde{O}(n + \sqrt{n}/\delta)$ time.

Proof

Set $\delta_0 := \delta / \lceil \log_2 n \rceil$. Build a balanced binary tree with n leaf nodes representing the items and $\lceil \log_2 n \rceil$ levels. At each node apply Corollary 4.4 thus obtaining a $(1-\delta)$ approximation of $\mathcal{S}(X)$ at the root node.

Lemma 4.6

Problem 2

Assume $\varepsilon \in (0, 1/2)$ and $1/\varepsilon \in \mathbb{N}^+$. Given a set X of n *distinct* integers in the interval $[1/\varepsilon, 2/\varepsilon)$, compute a set $A \subset \mathbb{N}$ that n-additively approximates $\mathcal{S}(X)$.

Lemma 4.6

We can solve Problem 2 in $\tilde{O}\left(n + \min\{\varepsilon^{-1} n^{1/2},\ \varepsilon^{-1} + \varepsilon^{-2}/n^{3/2}\}\right)$ time, which is at most $\tilde{O}(n + 1/\varepsilon^{5/4})$.

Depending on the parameters:

- When $n \leq \tilde{O}(1/\varepsilon^{1/2})$ run Algorithm 1
- Otherwise run Algorithm 2

Lemma 4.6 Algorithm 1

Algorithm 1

Directly apply Lemma 4.5 with $\delta := \varepsilon$, in $\tilde{O}(n + \sqrt{n}/\varepsilon)$ time. Because $n \leq \tilde{O}(1/\varepsilon^{1/2})$, the running time of Algorithm 1 is $\tilde{O}(n + 1/\varepsilon^{5/4})$.

Lemma 4.6

Lemma 2.7 [BW21]

Given n distinct positive integers $X = \{x_1, \dots, x_n\} \subseteq [\ell, 2\ell]$, there exists $\lambda = \tilde{\Theta}(\ell^2/n)$ such that, if $\lambda \leq \Sigma(X)/2$, then in $\tilde{O}(n)$ time we can construct a deterministic data structure supporting the following query in $O(1)$ time: given L, R such that $\lambda \leq L \leq R \leq \Sigma(X)/2$, report whether there exists $t \in [L, R]$ such that $t \in \mathcal{S}(X)$.

Lemma 4.6 Algorithm 2

Algorithm 2

Let $\sigma = \Sigma(X)$, and let λ be the threshold value from Lemma 2.7

- Initialize $A = \emptyset$
- Apply Lemma 4.5 in to compute a set A_δ that n-additive approximates $\mathcal{S}(X)$ up to λ.
- Insert all elements in $A_\delta \cap [0, \lambda]$ to A.
- Using the data structure from Lemma 2.7, compute an n-additive approximation of $\mathcal{S}(X) \cap [\lambda, \sigma/2]$ using binary search and insert it into A
- Let $A' := \{\sigma - a - n : a \in A\}$, $A \cup A'$ is an n-additive approximation of $\mathcal{S}(X)$ (up to σ)

Known reductions from the Partition Problem

Lemma C.1

One may assume w.l.o.g. that for any Subset Sum instance $\mathrm{OPT} \geq t/2$. Otherwise the instance can be solved exactly in $\tilde{O}(n)$ time.

Lemma C.2

Let $X \subset \mathbb{N}^+$ be a multiset with sum of elements $\sigma = \Sigma(X)$, and let $\varepsilon \in (0, 1/2)$. Given a set $A \subset \mathbb{N}$ that $\varepsilon\sigma/4$-additively approximates $\mathcal{S}(X)$, one can immediately solve $(1-\varepsilon)$-approximation Partition on X.

Proof of Lemma C.2

$t = \sigma/2$, and $\mathrm{OPT} = \max\{\Sigma(Y) : \Sigma(Y) \leq t, Y \subseteq X\}$. Given A, let $a := \max\{a \in A : a \leq t\}$. Claim:

$$(1-\varepsilon)\mathrm{OPT} \leq \min\{a, t(1-\varepsilon/2)\} \leq \mathrm{OPT},$$

Known reductions from the Partition Problem

Lemma C.3 MWW19

Given a multiset S of n integers from $[t]$, one can compute a multiset T in $O(n \log n)$ time such that:

- $\mathcal{S}(S; t) = \mathcal{S}(T; t)$.
- $|T| \leq |S|$.
- No element in T has multiplicity exceeding two.
- For every $y \in T$, there is a corresponding $x \in S$ such that $y = 2^k \cdot x$ for some $k \in \mathbb{N}$.

Lemma 4.1

If for some $c \geq 1$, Problem 2 can be solved in $\tilde{O}(n + 1/\varepsilon^c)$ time, then $(1 - \varepsilon)$-approximating *Partition* can also be solved in $\tilde{O}(n + 1/\varepsilon^c)$ time.