3SUM with Additive Combinatorics (C.-Lewenstein '15)

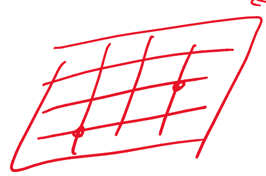
Q is $O(n^{2-\delta})$ possible?
special cases?
(e.g. for input in $[n^{2-\delta}]$, can use FFT)

Warm-Up Problem: 3SUM with Preprocessed Universe

Given sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of n integers,
preprocess s.t.
given any $A \subseteq \mathcal{A}$, $B \subseteq \mathcal{B}$, $C \subseteq \mathcal{C}$
can decide if $\exists a \in A, b \in B, c \in C$
$a+b=c$.

Intuition - "usually" not difficult
for most c, only small # ways to
write $c = a+b$.

- but bad exs exist
e.g. $\mathcal{A}, \mathcal{B}, \mathcal{C} = \{1, 2, 3, \ldots, n\}$
or arithmetic progression
or lattice



- but in these bad exs,
$|\mathcal{A}+\mathcal{B}|$ is linear instead of quadratic

Lemma (Cole-Hariharan '02)
For any sets A, B of n ints,
can compute $A+B$ in $\tilde{O}(|A+B|)$ rand. time
(output-sensitive)

Pf omitted. (by hashing + FFT). □

"BSG Thm" (Balog-Szemerédi-Gowers Thm)

"BSG Thm" (Balog-Szemerédi-Gowers Thm)

Given $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of size n.

Suppose $|\{(a,b) \in A \times B : a+b \in \mathcal{C}\}| \geq \alpha n^2$.

Then can extract subsets $A' \subseteq \mathcal{A}$, $B' \subseteq \mathcal{B}$ s.t.

(i) $|A' + B'| \leq O((1/\alpha)^5 n)$

(ii) $|A'|, |B'| \geq \Omega(\alpha n)$.

History: Balog, Szemerédi '94 complicated pf

Gowers '01

Balog '07 ← Tao's book)

Stronger version: "BSG Decomposition"

Given $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of size n.

Then $\exists$ subsets $A_1, \ldots, A_k \subseteq \mathcal{A}$, $B_1, \ldots, B_k \subseteq \mathcal{B}$ s.t.

(i) $A \times B = \bigcup_{i=1}^{k} (A_i \times B_i) \cup R$ ← small "remainder" set of size $\leq \alpha n^2$

(ii) $|A_i + B_i| = O((1/\alpha)^5 n)$

(iii) $k = O(1/\alpha)$.

Sol'n to Preproc-Univ. 3SUM:

Preprocessing: construct BSG decomposition in $\tilde{O}(n^2)$ rand. time (C.-Lewenstein '15)

Now given $A \subseteq \mathcal{A}$, $B \subseteq \mathcal{B}$, $C \subseteq \mathcal{C}$.

(i) for each $(a,b) \in R$,
check if $a \in A$, $b \in B$, $a+b \in C$.
$\Rightarrow O(\alpha n^2)$ time

(ii) for each k, compute
$(A_i \cap A) + (B_i \cap B)$ by Cole-Hariharan
in $\tilde{O}((1/\alpha)^5 n)$ time

$(A_i \cap A) + (B_i \cap B)$ by Cole-Hariharan in $\tilde{O}((1/\alpha)^5 n)$ time

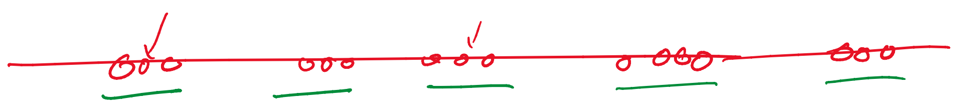
& check each $c \in C$

$\Rightarrow$ Time $\tilde{O}(\alpha n^2 + (\frac{1}{\alpha})^6 n)$

Set $\alpha = \frac{1}{n^{1/7}}$ $\Rightarrow$ $\boxed{\tilde{O}(n^{13/7})} = O(n^{1.86})$

Appl'n: 3SUM in "Clustered Case"

Suppose A, B, C are contained in m intervals of length ℓ.



For each element x, write $x = x^* + x^{**}$

where x^* = interval left endpt

$x^{**} \in [\ell]$.

Preprocess $\{a^* : a \in A\}$, $\{b^* : b \in B\}$, $\{c^* : c \in C\}$ in $\tilde{O}(m^2)$ time

For each $i, j \in [\ell]$,

solve 3SUM for
$\{a^* : a^* + i \in A\}$,
$\{b^* : b^* + j \in B\}$
$\{c^* : c^* + i + j \in C\}$

or $\{c^* - \ell : c^* + i + j - \ell \in C\}$

$\Rightarrow$ total time $\boxed{\tilde{O}(m^2 + \ell^2 m^{13/7})}$

Appl'n: Offline Jumbled Text Indexing

$t \in [\sigma]^*$

Appln: Offline Jumbled ...

text $t_1 \cdots t_n \in [\sigma]^*$

given query vectors $c_1, \ldots, c_n \in [n]^\sigma$,
for each c_k, decide if $\exists$ substring $t_{i+1} \cdots t_j$ with frequency vector c_k.

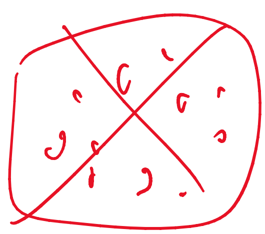
Let a_i = frequency vector for $t_1 \cdots t_i$

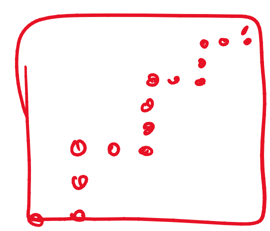
For each c_k, decide iff $\exists i, j$, $a_j - a_i = c_k$.

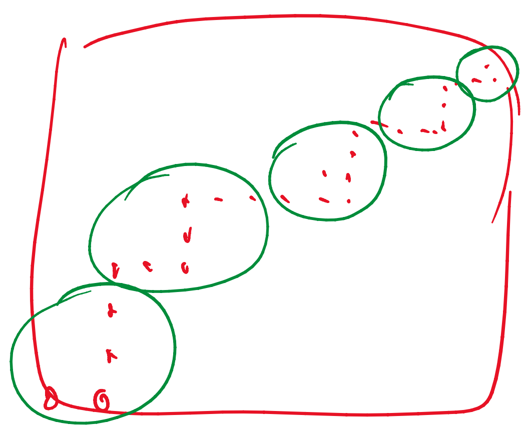
3SUM for vectors in σ dims.

but input satisfies bounded-diff property

$(a_{i+1} - a_i \in \{0,1\}^\sigma)$







divide set of vectors in $\frac{n}{\ell}$ clusters
each is in $[\ell]^\sigma$.

$$\Rightarrow \tilde{O}\left(\left(\frac{n}{\ell}\right)^2 + (\ell^\sigma)^2 \left(\frac{n}{\ell}\right)^{\frac{13}{7}} \right)$$

$$= \tilde{O}\left(\frac{n^2}{\ell^2} + \ell^{2\sigma - \frac{13}{7}} n^{13/7} \right)$$

$$= \tilde{O}\left(\frac{n^2}{\ell^2} + \ell^{2\sigma - \frac{13}{7}} n^{13/7} \right)$$

Set $\ell^{2\sigma - \frac{13}{7} + 2} = n^{2 - 13/7}$, $\ell^{2\sigma + \frac{1}{7}} = n^{\frac{1}{7}}$

$\ell = n^{\frac{1}{14\sigma + 1}}$

$$\Rightarrow \boxed{\tilde{O}\left(n^{2 - \frac{2}{14\sigma + 1}} \right)}$$

(improves to $\tilde{O}\left(n^{2 - \frac{2}{\sigma + 13}}\right)$)

e.g. $\sigma = 2$: $O(n^{1.94})$

(improves to $O(n^{1.859})$)

(Compare with conditional LB $\Omega\left(n^{2 - \frac{4}{\sigma} - \delta}\right)$)

Other appl'ns - bdd-diff (min,+)-convol in $O(n^{1.859})$ time