BETTER ANALYSIS of SET-FIND

Def. A set $W \subseteq \mathbb{R}^d$ of n points is (t, γ, β)-stretched at scale ℓ if for at least γ fraction of directions u there is a (partial) matching $M_u \subseteq W \times W$ with $\frac{\beta n}{4}$ disjoint pairs (v_i, w_i) s.t.

$$\text{t-stretched pair} \begin{cases} \text{i) } |v_i - w_i| \leq \ell \\ \text{ii) } \langle u | v_i - w_i \rangle \geq \frac{t \cdot \ell}{\sqrt{d}} \end{cases}.$$

Theorem 5: For any $\gamma, \beta > 0$ there exists $C = C(\gamma, \beta)$ s.t. • if a unit ℓ^2-representation is (t, γ, β)-stretched at scale ℓ, for some $\ell > 0$, then $t \leq C \cdot (\log n)^{1/3}$.

Cor. For any $\gamma > 0$, the probability that SET-FIND with parameters $c' = \frac{b}{8}$, $\sigma = \frac{b^2}{48}$, $\Delta = \frac{\sigma^2}{C^2 \cdot (\log n)^{2/3}}$ where $b = 0{,}878 \cdot c \cdot (1-c)$, removes in step 4 a matching of size $> c' \cdot n$, is at most γ.

$C = C(\gamma, 4c')$

Proof of Cor: every deleted pair v_i, v_j satisfies

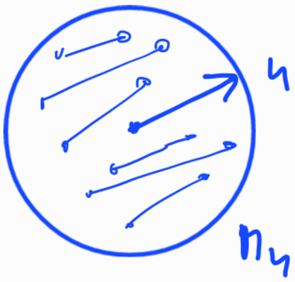
i) $(v_i - v_j)^2 \leq \Delta$ .... i.e., $|v_i - v_j| \leq \frac{\sigma}{C (\log n)^{1/3}}$

ii) $\langle u | v_i - v_j \rangle \geq \frac{2\sigma}{\sqrt{n}} = 2 \cdot C \cdot (\log n)^{1/3} \frac{\sigma}{C (\log n)^{1/3} \cdot \sqrt{n}}$

If the probability is larger than γ, then for at least γ fraction of directions there is a matching with $> c' n$ $(2 \cdot C \cdot (\log n)^{1/3})$-stretched pairs, that is, the SDP solution $v_1, \dots, v_n$ is $(2C (\log n)^{1/3}, \gamma, 4c')$-stretched at scale $\sqrt{\Delta}$ - a contradiction with THM 5. ☒

Def B: A set $W \subseteq \mathbb{R}^d$ of n points is (ε, δ)-matching-covered at scale ℓ if for every unit vector $u \in \mathbb{R}^d$, there exists a (partial) matching $M_u \subseteq W \times W$ s.t.

1. $\forall (v,w) \in M_u$, i) $\|v-w\| \leq \ell$ short
   large projection ii) $\langle u | v-w \rangle \geq \varepsilon$
2. $\forall v \in W$, $\mu(u : v \text{ matched in } M_u) \geq \delta$.



The set of these matchings M_u - matching cover.

Note: main difference between Def. A and B: in B, every point participates in M_u with prob. $\geq \delta$, for a random direction.

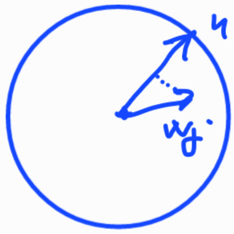
Lemma 6: If a set W of n points is (t, γ, β)-stretched at some scale ℓ, then it contains a subset X of $\Omega(\gamma \cdot \beta \cdot n)$ points that is (ε, δ)-matching covered at scale ℓ, where $\delta = \frac{\beta \cdot \gamma}{4}$, $\varepsilon = \frac{t \cdot \ell}{\sqrt{d}}$.

Proof: skipped - counting

Let H denote the multigraph on X (from Lemma 6) whose set of edges is the union of all matchings M_u in the matching cover.

For $v_i \in X$, let $\text{Ball}(v_i, r) = \{v_j \in X : \text{distance}(v_i, v_j) \text{ in } H \text{ at most } r\}$

Def C: A set $\{w_1, w_2, \dots\} \subseteq \mathbb{R}^d$ of vectors is an (ε, δ)-cover if for at least δ-fraction of unitvectors $u \in \mathbb{R}^d$, there exists j s.t. $\langle u, w_j \rangle \geq \varepsilon$.



"every unit u is close to some w_j"

For any fixed $v_i \in X$, $\{v_j - v_i \mid v_j \in \mathrm{Ball}(v_i, 1)\}$ is (ε, δ)-cover.

Proof: Every v_i is matched in at least δ directions u, i.e., for each such direction u there is $v_j \in \mathrm{Ball}(v_i, 1)$, and each pair (v_i, v_j) in the matching has the property $\langle u | v_i - v_j \rangle \geq \varepsilon$. $\square$

Notation: Let

$$S_r = \{v_i \in X \mid \{v_j - v_i \mid v_j \in \mathrm{Ball}(v_i, r)\} \text{ is an } (\tfrac{\varepsilon r}{2}, 1 - \tfrac{\delta}{2})\text{-cover}\}$$

Note: at this point, it is not clear that $S_1 \neq \emptyset$.

Lemma 7: ← the core of the analysis (measure of concentration)

i) $S_1 = X$

ii) There exist constants $\eta = \eta(\delta)$, $\rho = \rho(\delta)$ s.t. for $r \leq \eta \cdot t$,

$$|S_{r+1}| \geq \rho |S_r|.$$

Proof: postponed / skipped

Proof of THM 5: Assume W is (t,η,β)-stretched.

By Lem. 6, there is $X \subseteq W$ that is (ε,δ)-matching covered,
for $\delta = \Omega(\eta)$, $\varepsilon = \frac{t \cdot \ell}{\sqrt{d}}$, and each edge in matching has squared lenght at most ℓ^2.

Assume $t = \Omega(\log^{1/3} n)$ (otherwise THM 5 already holds).

By Lem 7, for $r = \eta \cdot t$, we get $\quad r = o(\log n)$

$$|S_r| \geq \rho^{r-1} \cdot |S_1| = \Omega(\rho^{r-1} \cdot n) \gg 1 \quad \text{i.e., } S_r \neq \emptyset$$

$\forall v_i \in S_r : \{v_j - v_i \mid v_j \in \mathrm{Ball}(v_i, r)\}$ is an $(\frac{\varepsilon r}{2}, 1 - \frac{\delta}{2})$-cover

i.e., for at least $1 - \frac{\delta}{2}$ directions, there exists

$\underbrace{v_j \in \mathrm{Ball}(v_i, r)}_{}$ s.t. $\underline{|\langle v_i - v_j \mid u \rangle| \geq \frac{\varepsilon r}{2}}$.

by Δ-inequal: $|v_i - v_j|^2 \leq r \cdot \ell^2$,

i.e., $\underline{|v_i - v_j| \leq \sqrt{r} \cdot \ell}$.

stretch of $v_i - v_j$

$$\frac{\frac{\varepsilon r}{2}}{\frac{\sqrt{r} \cdot \ell}{\sqrt{d}}} \overset{\varepsilon = \dots}{=} \frac{\frac{t \cdot \ell \cdot \sqrt{r}}{\sqrt{d} \cdot 2}}{\frac{\ell}{\sqrt{d}}} = \frac{t \cdot \sqrt{r}}{2} \overset{r = \dots}{=} \frac{\sqrt{\eta} \cdot t^{3/2}}{2} = \Omega(t^{3/2}),$$

with prob. $\geq 1 - \frac{\delta}{2}$

By Aux. lemma: Prob. that stretch is $\geq 2\sqrt{2 \log n}$ is $\leq e^{-2 \log n} \leq \frac{1}{n^2}$

$$\Rightarrow \quad t^{3/2} = O(\log^{1/2} n) \quad \Rightarrow \quad t = O(\log^{1/3} n). \quad \blacksquare$$