Def: $\mathcal{U}$ is a $T(\alpha)$-database update algorithm for $\mathcal{A}$ if $\forall D \in \mathcal{D}$, every $(\mathcal{U}, D, Q, \alpha, L)$-database update sequence has $L \leq T(\alpha)$

Def: $(F(\varepsilon), \gamma)$-private distinguisher: ε-DP alg. that, given $D \in \mathcal{D}$, $S \subseteq \mathcal{S}$, outputs $f^* \in Q$ s.t.

$$|f^*(D) - f^*(S)| \geq \max_{f \in Q} |f(D) - f(S)| - F(\varepsilon) \text{ w.p. } \geq 1 - \gamma$$

$\mathcal{I}(D, \alpha, \varepsilon_0)$

- Init $D^0 \in \mathcal{S}$
- For $t = 1$ to $T\left(\frac{\alpha}{2}\right)$:
  - Let $f^t = \text{Distinguish}(D, D^{t-1})$
  - Let $\hat{v}^t = f^t(D) + \text{Lap}\left(\frac{1}{|D|\varepsilon_0}\right)$
  - if $|\hat{v}^t - f^t(D^{t-1})| < \frac{3\alpha}{4}$, output D^{t-1}
  - else $D^t = \mathcal{U}(D^{t-1}, f^t, v^t)$
- Output $D^{T(\alpha/2)}$

Thm: ε-DP if $\varepsilon_0 \le \frac{\varepsilon}{2T(\alpha/2)}$

(ε, δ)-DP if $\varepsilon_0 \le \frac{\varepsilon}{4\sqrt{T(\frac{\alpha}{2})\ln\frac{1}{\delta}}}$

Pf: Basic/advanced composition: $2T(\frac{\alpha}{2})$ ε_0-DP steps (Distinguish, $\hat{v}^t$).

Thm: Given $(F(\varepsilon), \gamma)$-private distinguisher, $T(\alpha)$-database update algorithm. Then w.l. prob. $\ge 1-\beta$, $\mathcal{I}$ returns $S \in \mathcal{S}$ s.t.

$$\max_{f \in Q} |f(D) - f(S)| \le \alpha \text{ for any}$$

$$\alpha \ge \max\left(\frac{8 \log\left(\frac{2T(\alpha/2)}{\beta}\right)}{\varepsilon |D|}, 8F(\varepsilon_0) \right)$$

as long as $\gamma \le \frac{\beta}{2T(\alpha/2)}$

Pf: w.l. prob. $\ge 1-\beta/2$, $|\hat{v}^t - f^t(D)| \le \frac{\alpha}{8}$ $\forall t$ (Laplace)

w.l. prob. $\ge 1-\beta/2$, $|f^t(D) - f^t(D^{t-1})| \ge \max_{f \in Q} |f(D) - f^t(D^{t-1})| - \frac{\alpha}{8}$ $\forall t$ (private distinguisher)

If return $D^{T(\alpha/2)}$, then check

$|\hat{v}^t - f^t(D^{t-1})| \geq \frac{3\alpha}{4}$ and

$|\hat{v}^t - f^t(D)| \leq \frac{\alpha}{8}$

$\Rightarrow |f^t(D) - f^t(D^{t-1})| \geq \frac{3\alpha}{4} - \frac{\alpha}{8} = \frac{5\alpha}{8} \geq \frac{\alpha}{2}$

$\Rightarrow$ so $T(\frac{\alpha}{2})$ distinguishes queries w/ param. $\frac{\alpha}{2}$

$\Rightarrow \frac{\alpha}{2}$-accurate $\forall f \in Q$.

Else return D^{t-1} for some $t < T(\alpha/2)$

$\Rightarrow |\hat{v}^t - f^t(D^{t-1})| < \frac{3\alpha}{4}$

$\Rightarrow |f^t(D) - f^t(D^{t-1})| \leq |\hat{v}^t - f^t(D^{t-1})| + |f^t(D) - \hat{v}^t|$

$\leq \frac{3\alpha}{4} + \frac{\alpha}{8} = \frac{7\alpha}{8}$

$\Rightarrow$ by property of private distinguisher,

$$\max_{f \in Q} |f(x) - f(D^{t-1})| \leq |f^t(D) - f^t(D^{t-1})| + F(\varepsilon_0)$$

$$\leq \frac{7\alpha}{8} + F(\varepsilon_0) \leq \alpha \quad \checkmark$$

use EM as private distinguisher: domain Q,

quality score $q(D, f) = |f(D) - f(D^t)|$

Thm: $\mathcal{M}$ is an $(F(\varepsilon), \gamma)$-private distinguisher for

$$F(\varepsilon) = \frac{2}{|D|\varepsilon} \log \frac{|Q|}{\gamma}$$

Plug this thm into IC utility thm just correct ε_c:

Thm: Given $T(\alpha)$-database update algorithm, IC w/ $\mathcal{M}$ distinguisher is ε-DP, returns S s.t.

$$\max_{f \in Q} |f(D) - f(S)| \leq \alpha \text{ for } \text{w.p. } \geq 1-\beta.$$

$$\alpha = \frac{8\, T(\frac{\alpha}{2})}{|D|\varepsilon} \left(\log \frac{|Q|}{\gamma}\right)$$

is (ε, δ)-DP, returns S s.t. $\max_{f \in Q} |f(D) - f(S)| \leq \alpha$ for

$$\alpha = \frac{16\sqrt{T(\frac{\alpha}{2}) \log(1/\delta)}}{|D|\varepsilon} \log \frac{|Q|}{\gamma}$$

s. long as $\gamma \leq \beta / 2T(\frac{\alpha}{2})$

Application: Another data structure, database update alg allowing us to handle nonlinear queries.

- Combination of nets, incremental construction.

Def: Median data structure: S is a collection of databases: $S \subseteq \mathcal{D}$. For query $f: \mathcal{D} \to \mathbb{R}$, extend to $f(S) = \text{median}(\{f(D) : D \in S\})$

Median Mechanism update rule:

Init: $D_0 = N_\alpha(Q)$

$MM(D^t, f^t, v^t)$:

- If $v^t < f^t(D^t)$
  - output $D^{t+1} = D^t \setminus \{D \in D^t : f_t(D) > f_t(D^t)\}$
- Else ($v^t > f^t(D^t)$)
  - output $D^{t+1} = D^t \setminus \{D \in D^t : f_t(D) < f_t(D^t)\}$

Thm: For any Q, MM is a $T(\alpha) = \log |N_\alpha(Q)|$ database update algorithm

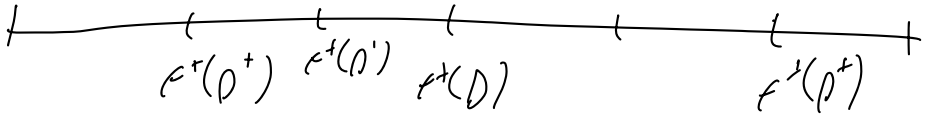
Pf: Need to show update sequence longer than $\log |N_\alpha(Q)|$.

Let $D' \in N_\alpha(Q)$ s.t. $\max_{f \in Q} |f(D) - f(D')| \leq \alpha$ (since α-net)

By def of update sequence, know

1) $|f^t(D') - f^t(D)| > \alpha$, $f^t(D') = \text{median}\{f^t(\hat{D}) : \hat{D} \in D^t\}$

2) $|f^t(D) - v^t| < \alpha$



$\Rightarrow f^t(D')$ and $f^t(D)$ on same side of $f^t(D^t)$

$\Rightarrow$ MM never removes D' from data structure

$\Rightarrow |D^t| \geq 1 \quad \forall t$

By def of median, $|D^t| \leq \frac{1}{2}|D^{t-1}|$

$\Rightarrow$ after L update steps, $1 \leq |D^L| \leq \frac{1}{2^L}|N_\alpha(Q)|$

$\Rightarrow L \leq \log |N_\alpha(Q)|$

Plug into (ε, δ)-version of IC with trivial metric: all databases of size $|D| = n$!

$\Rightarrow$ want $\forall \alpha$ size true database is not (in public).

Thm: IC using MM as database update rule with $N_\alpha(Q) = \{\text{all databases of size } n\}$, exp. mech. $\Rightarrow$ private distinguisher, is (ε, δ)-DP and w.prob. $\geq 1-\beta$ returns a data structure S s.t.

$$\max_{f \in \mathcal{Q}} |f(S) - f(D)| \leq \frac{16\sqrt{\log|\mathcal{X}| \log\frac{1}{\delta}}\ \log\left(\frac{2|\mathcal{Q}|n\log|\mathcal{X}|}{\beta}\right)}{\sqrt{n}\cdot\varepsilon}$$

for any set of $\frac{1}{n}$-sensitive queries $\mathcal{Q}$.

Pf: Previous thm w/ abstract $T(\mathcal{Q})$-update rule

$$\alpha = \frac{16\sqrt{T(\alpha/2)\log(1/\delta)}}{n\varepsilon}\log\frac{2|\mathcal{Q}|T(\alpha/2)}{\beta}$$

$$w(MW, T(\tfrac{\alpha}{2})) \leq \log(|\mathcal{X}|^n) = n\log|\mathcal{X}|$$

$$\leq \frac{16\sqrt{\log|\mathcal{X}|\log\frac{1}{\delta}}}{\sqrt{n}\,\varepsilon}\cdot\log\left(\frac{2|\mathcal{Q}|n\log|\mathcal{X}|}{\beta}\right) \quad \checkmark$$

Note: Doesn't work for pure DP: n is cancelled out!

Online Versions: Just use Numeric Sparse to determine if error is large!

$\Rightarrow$ same as private MW.

$\Rightarrow$ Basically no extra loss - just constants, logs.

$\Rightarrow$ online bounds = offline bounds!