Today:
- Recap MW guarantee
- Query release via Synthetic Data Distributions
- Online learning $\Rightarrow$ Query Release : MW-EM.

What will we get? k queries (linear)

m : size of the universe.

Gaussian/Laplace : ℓ_∞ error α when $n \geq n_{Gauss} = \frac{C_{\varepsilon,\delta}\sqrt{k}\log k}{\alpha}$

Projection : ℓ_2 error α when $n \geq n_{Proj} \approx \frac{C'_{\varepsilon,\delta}\sqrt{\log(m)}}{\alpha^2}$

MW-EM : ℓ_∞ error α — $n \geq n_{MWEM} = \frac{C''_{\varepsilon,\delta}(\log k)\sqrt{\log m}}{\alpha^2}$

Think of $\log m \approx$ data dimension

Thresholds : $k = m$. $n_{MWEM} \gtrsim \frac{\log^{3/2}(m)}{\alpha^2}$ vs. $n_{Gauss} \gtrsim \frac{\log^{3/2}(m)}{\alpha}$ with binary tree

"Pairwise marginals" $\mathcal{U} = \{0,1\}^d$ $n_{MWEM} \gtrsim \frac{\sqrt{d}\cdot\log(d)}{\alpha^2}$ vs. $\frac{d\sqrt{\log(d)}}{\alpha}$ For Gaussian

$\varphi_{ij}(x) = \begin{cases} 1 & \text{if } x_i = x_j = 1 \\ 0 & \text{o.w.} \end{cases}$

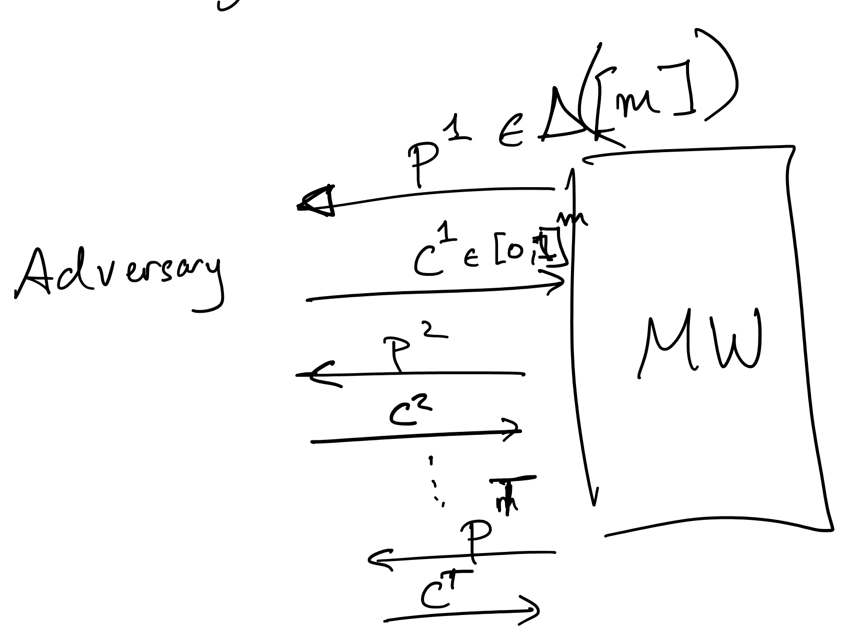
$m = 2^d$

$k = d^2$

MW outputs as vectors:

- Last time : MW takes a (random) action at each time.
- Today : MW outputs a vector $p^t \in \Delta([m])$ at each time

(notation switch.)



Theorem : $\forall$ adversary $\forall p^* \in \Delta([m])$

$$\frac{1}{T}\sum_{t=1}^{T}\langle c^t, p^t\rangle - \frac{1}{T}\sum_{t=1}^{T}\langle c^t, p^*\rangle \leq 2\sqrt{\frac{\ln(m)}{T}}$$

$\langle c^t, p^t\rangle = \mathbb{E}_{a\sim p^t}(c^t_a)$

with prob. 1 !

Proof: • Look at equation (10) ☺

• $\frac{1}{T}\sum_{t=1}^{T}\langle c^t, p^*\rangle = \mathbb{E}_{a\sim p^*}\left(\frac{1}{T}\sum_{t=1}^{T} c^t_a\right) \geq \min_{a^*} \frac{1}{T}\sum_{t=1}^{T} c^t_{a^*}$. ☒

Query Release Via Synthetic Data Distributions

→ Given $F = \{f_1, \dots, f_k\}$. $\qquad f_i(\vec{x}) = \frac{1}{n}\sum_{j=1}^{n} \varphi_i(x_j)$

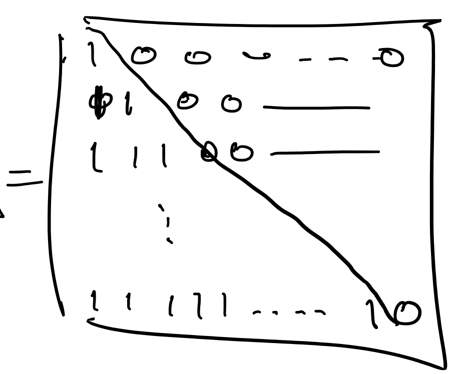
$\varphi_i : \mathcal{U} \to [0,1]$

→ Histogram $(\vec{h_x})_u = \frac{\#\{j : x_j = u\}}{n}$

→ Want $\vec{a} \approx F h_{\vec{x}}$

Thresholds: $\mathcal{U} = \{1, \dots, m\}$

$\varphi_i(x) = \begin{cases} 1 & \text{if } x \le i \\ 0 & \text{o.w.} \end{cases}$

$F_{\text{thres}} =$ 

Idea #1: Release a distribution $\hat{p}$ on $\mathcal{U} = \{1, \dots, m\}$.

s.t. $\text{error}(\hat{p}) = \| F\hat{p} - Fh_x \|_\infty \le \alpha$

↳ $\forall i = 1, \dots, k$: $\left| \mathbb{E}_{a \sim \hat{p}} \varphi_i(\hat{p}) - \frac{1}{n}\sum_{j=1}^{n} \varphi_i(x_j) \right| \le \alpha.$

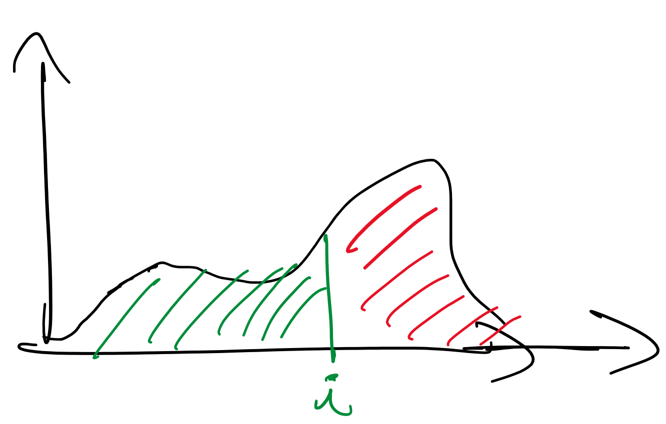
$|\langle \varphi_i, \hat{p} \rangle - \langle \varphi_i, h_x \rangle| \le \alpha$

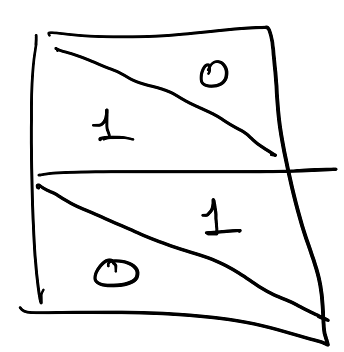
– Absolute values ☹

– Trick: consider sets of queries closed under complements.

If $\varphi_i \in F \implies \varphi_i' = 1 - \varphi_i$ is also in F.

For thresholds, add $\varphi_i'(x) = \begin{cases} 0 & \text{if } x \le i \\ 1 & \text{if } x > i. \end{cases}$



$F_{\text{thres}} =$ 

– $|y| = \max(y, -y)$.

$|\langle \varphi_i, \hat{p} - h_x \rangle| = \max\left(\langle \varphi_i, \hat{p} - h_x \rangle, \langle -\varphi_i, \hat{p} - h_x \rangle \right)$

entries add to 0

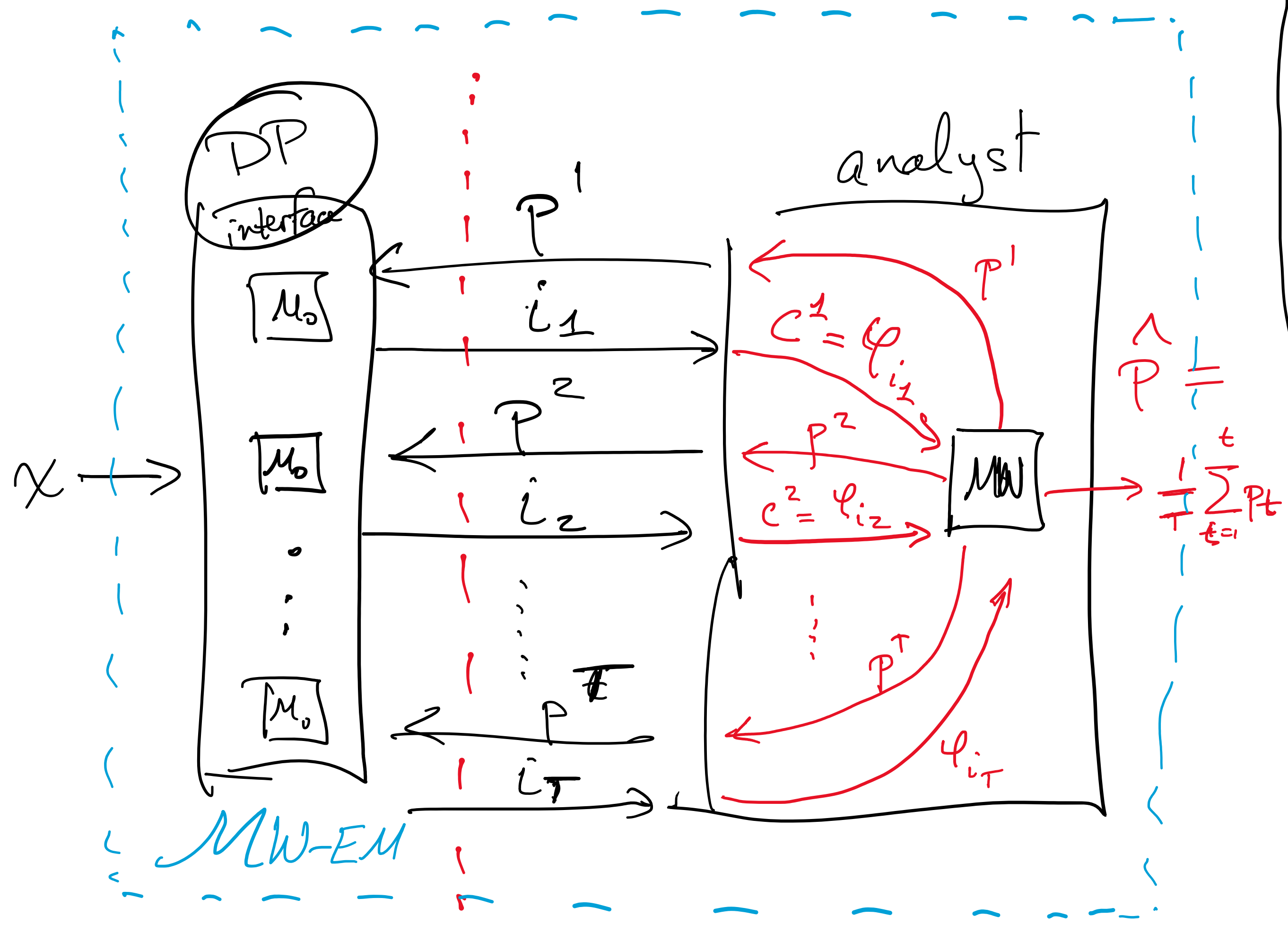
$= \max\left(\langle \varphi_i, \hat{p} - h_x \rangle, \langle \vec{1} - \varphi_i, \hat{p} - h_x \rangle \right)$

If F is closed under comp, then

$\therefore \ \text{error}(\hat{p}) = \max_{i=1,\dots,k} \langle \varphi_i, \hat{p} - h_x \rangle.$

From Online Learning to Query Release

Goal: Design $\mathcal{M}$: $x \longrightarrow \boxed{\mathcal{M}} \longrightarrow \hat{p}$ s.t. $\max_{i=1,\dots,k} \langle \varphi_i, \hat{p} - h_x \rangle \leq \alpha$ with high prob.



Design $\mathcal{M}$ as an interaction between:

— analyst who propose sequence $p^1, p^2, \dots, p^T \in \Delta([m])$

— DP "interface" run T executions of exp. mech.

$\mathcal{M}_0$:

- Goal: Find i s.t. $\langle \varphi_i, p^t - h_x \rangle \approx \text{error}(p^t)$
- Input: x, ε_0, p^t
- Output: i. $\Pr(M_0 = i) \propto e^{\varepsilon_0 n \cdot \text{score}(i; x)/2}$

$\text{score}(i; x) = \langle \varphi_i, \hat{p} - h_x \rangle$

Observer: $\text{error}(p^t) = \max_i \text{score}(i; x)$

- Sensitivity of score? If x, y neighbors

$\vec{x}, \vec{y}$ differ in one record, $x_{old} \to x_{new}$.

$$|\langle \varphi_i, p^t - h_x \rangle - \langle \varphi_i, p^t - h_y \rangle| = |\langle \varphi_i, h_x - h_y \rangle| = \left|\tfrac{1}{n}(\varphi_i(x_{old}) - \varphi_i(x_{new}))\right| \leq 1/n$$

MW-EM$(x, F, \varepsilon, \delta)$

$p^1 \leftarrow (1, \dots, 1)$ ~ length m

for $t = 1$ to T:

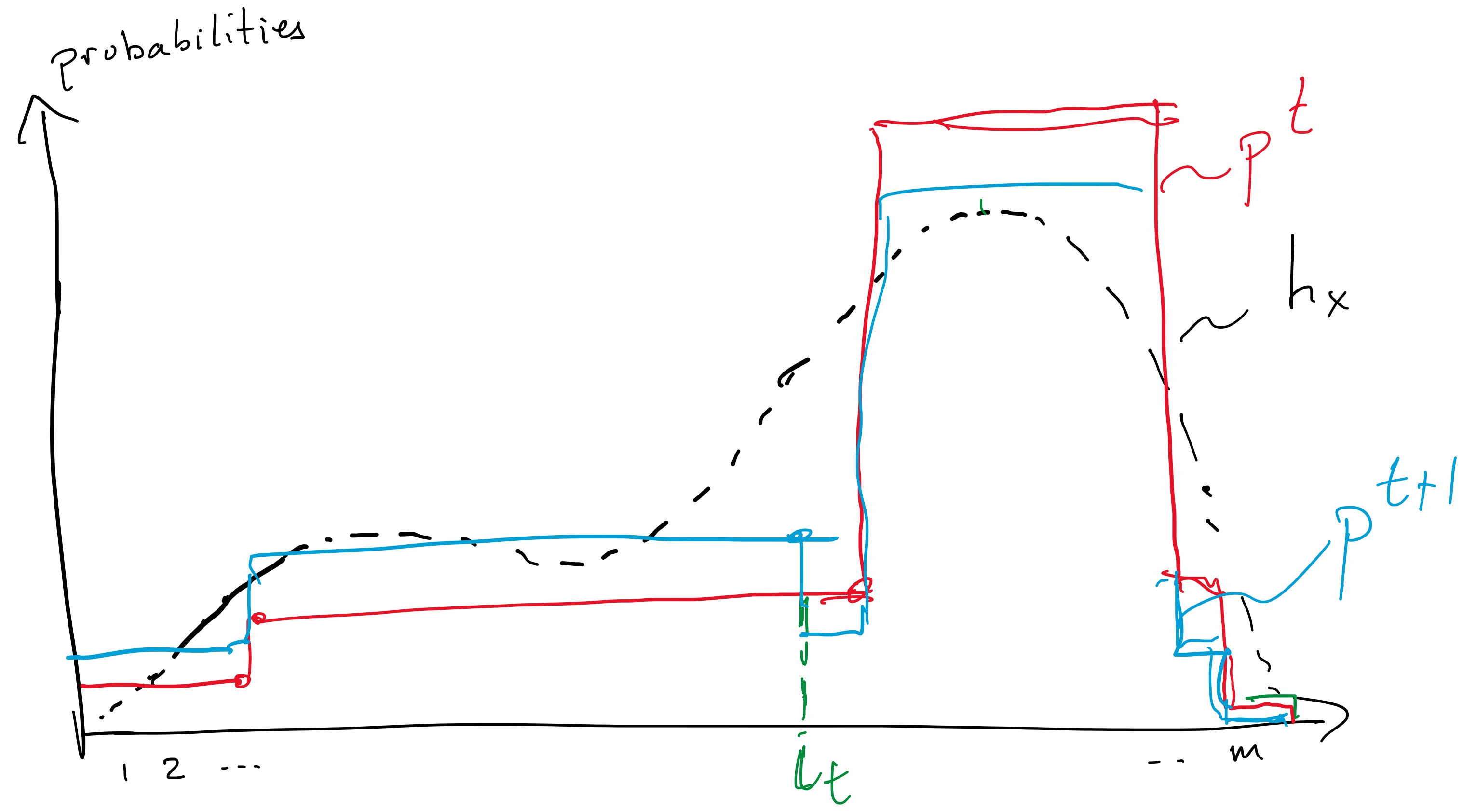
- $i_t \leftarrow \mathcal{M}_0(x, \varepsilon_0, p^t)$
- $c^t \leftarrow \varphi_{i_t}$
- $p^{t+1} \leftarrow \text{MW-Update}(p^t, c^t, \eta)$ ← $\sqrt{\frac{\ln(m)}{T}}$

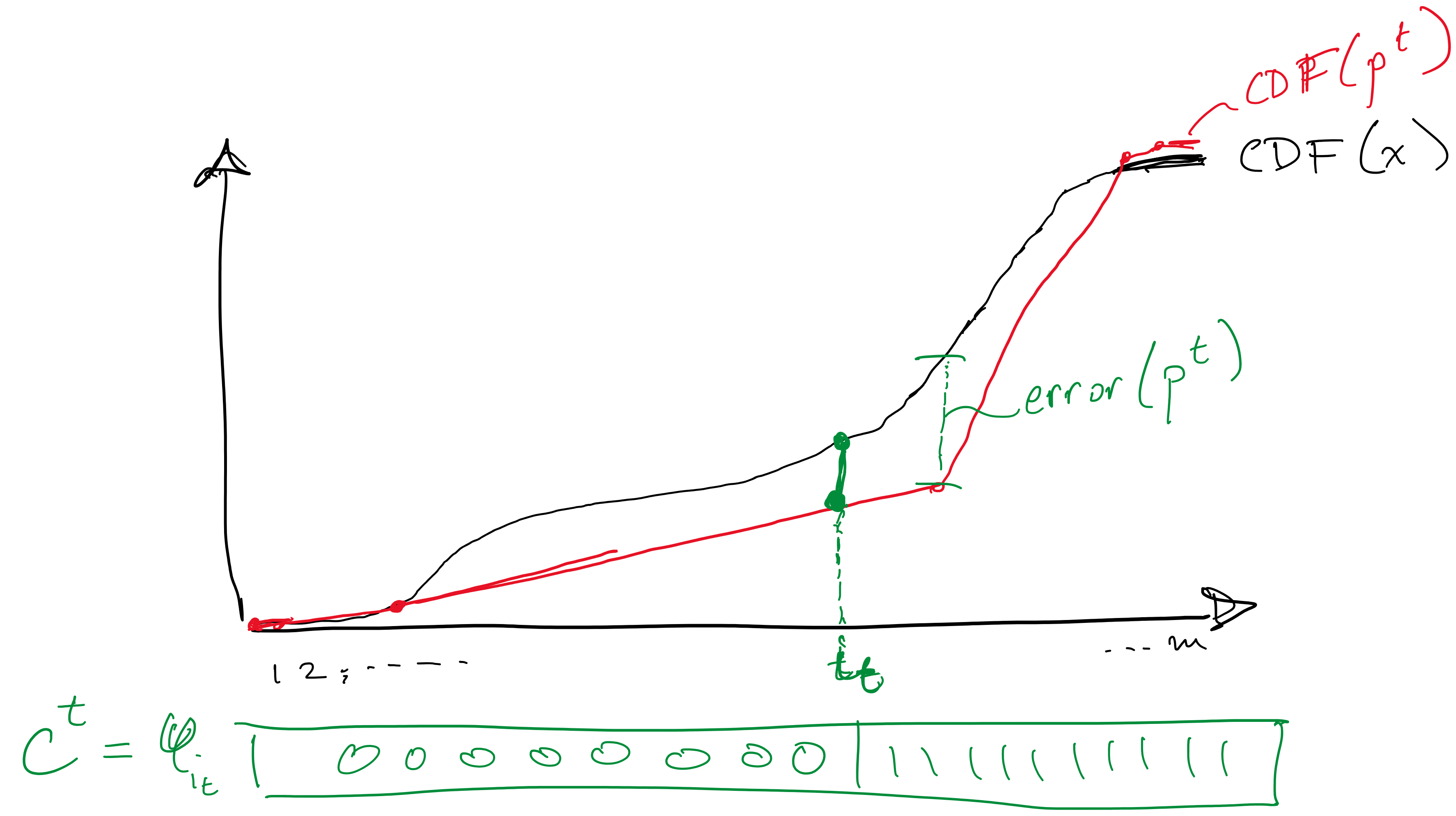
Return $\frac{1}{T}\sum_{t=1}^{T} p^t$.

Why is this not crazy?

- At each stage, p^t gets "closer" to h_x
- We will use MW as a black box.
  (but can show that "KL divergence" $D(h_x \| p^t)$ decreases).

Example: Thresholds





How does the analysis work?

- Exp. Mech i_t s.t. $\langle \varphi_{i_t}, p^t - h_x \rangle \approx \text{error}(p^t)$
- Suppose error (p^t) is high at all T time steps. $\gtrsim \alpha$

$$\underset{\text{relative to } h_x}{\text{Regret}} = \frac{1}{T}\sum_t \langle c^t, p^t \rangle - \frac{1}{T}\sum_t \langle c^t, h_x \rangle$$

$$= \frac{1}{T}\sum_t \langle c^t, p^t - h_x \rangle \gtrsim \frac{1}{T}\sum_t \alpha = \alpha.$$

- Theorem $\Rightarrow$ $\alpha \leq \text{Regret} \leq 2\sqrt{\frac{\ln(m)}{T}}$

So ...? $T \leq \ln(m)/\alpha^2$ ....

when T is large, average error will be small ☺

Privacy: MWEM is composition of T alg's, each $(\varepsilon_0, 0)$-DP.

$\therefore$ (ε, δ)-DP for

$$\varepsilon = \varepsilon_0 \sqrt{2T \ln(1/\delta)} + T \cdot \varepsilon_0 \cdot \frac{e^{\varepsilon_0} - 1}{e^{\varepsilon_0} + 1} = \Theta\left(\varepsilon_0 \sqrt{T \ln(1/\delta)}\right)$$

← Lecture 9 ?

(when $\varepsilon < 1$)

$$\therefore \varepsilon_0 \approx \frac{\varepsilon}{\sqrt{T \ln(1/\delta)}} = \frac{1}{c_{\varepsilon,\delta} \sqrt{T}}.$$

Accuracy?

Lecture 6 $\Rightarrow$ With prob $\geq 1 - \beta/T$, each execution of Exp. Mech.

$$\langle \varphi_i, p^t - h_x \rangle \geq \text{error}(p^t) - \underbrace{\frac{4 \ln(k \cdot T/\beta)}{\varepsilon_0 n}}_{\alpha_0}$$

↑ max. score

$$\text{error}(\hat{p}) = \max_i \langle \varphi_i, \hat{p} - h_x \rangle$$

← convex function of $\hat{p}$ because it's a max of linear functions.

$$\leq \frac{1}{T} \sum_{t=1}^{T} \underbrace{\max_i \langle \varphi_i, p^t - h_x \rangle}_{\text{error}(p^t)}$$

← by Jensen's ineq.

with prob $\geq 1-\beta$:

$$\leq \frac{1}{T} \sum_{t=1}^{T} \left(\langle c^t, p^t - h_x \rangle + \alpha_0 \right)$$

$$= \underbrace{\frac{1}{T} \sum_t \langle c^t, p^t \rangle - \frac{1}{T} \sum_t \langle c^t, h_x \rangle}_{\text{regret}} + \alpha_0$$

$$\leq 2 \sqrt{\frac{\ln(m)}{T}} + \frac{4 \ln(k \cdot T/\beta)}{n \varepsilon_0} \cdot c_{\varepsilon,\delta} \cdot \sqrt{T}.$$

$$T = \frac{\ln(m) \cdot \sqrt{n}}{\sqrt{\ln(k T/\beta)}\, c_{\varepsilon\delta}}$$

$$\alpha = \tilde{O}\left(\frac{c_{\varepsilon,\delta}^{1/2} \ln^{1/4}(m) \cdot \ln^{1/2}(k/\beta)}{\sqrt{n}} \right)$$

$\therefore$ $\text{Error}(\hat{p}) \leq \alpha$ w.p. $\geq 1-\beta$ when $n \geq \tilde{O}\left(\frac{c_{\varepsilon,\delta} \ln(k/\beta) \sqrt{\ln(m)}}{\alpha^2} \right)$.

(dropping log log m).