Private MW for online query release.

$q_1, \cdots, q_k$ chosen adaptively. dataset $X \in \mathbb{N}^{|\mathcal{X}|}$

$p := \frac{X}{n}$ empirical distribution

our algorithm should answer

$a_1, a_2, \cdots, a_k$ with error $\leq \alpha$, $\text{error}(q_i) = |a_i - q_i^T p|$

and satisfy ε-DP

Question: Is there an algorithm with sample complexity $\text{poly}(\log k, \log|\mathcal{X}|, \frac{1}{\alpha}, \frac{1}{\varepsilon})$?

Algorithm. PMW. take ε, N, k as inputs, q_i stream Q

start with $\hat{p}_1 = \frac{1}{|\mathcal{X}|}$, $i = 1$

for $t = 1, \cdots, N$: $\hat{\alpha} = \alpha + \underbrace{\text{Lap}(\frac{4}{\varepsilon_0 n})}_{w_t} - \frac{\tau}{2}$

while $i \leq k$:

get q_i from Q

if $|q_i^T \hat{p}_t - q_i^T p| + \underbrace{\text{Lap}(\frac{8}{\varepsilon_0 n})}_{\nu_i} \geq \hat{\alpha}$

output $a_i = q_i^T \hat{p}_t$

$i = i+1$

else:

$a_i = q_i^T p + \underbrace{\text{Lap}(\frac{2}{\varepsilon_0 n})}_{z_t}$

update $\hat{p}_{t+1} = MW(\hat{p}_t, \ell_t)$

with $\ell_t = q_i \cdot \text{sign}(q_i^T \hat{p}_t - a_i)$

$t = t+1$

$i = i+1$

break (out of while loop)

for $i = 1, 2, \cdots, k$

$\varepsilon_0 = \frac{\varepsilon}{2\sqrt{N}}$

$\tau = \frac{8 \log(2Nk/\delta)}{\varepsilon_0 n}$

Theorem (Privacy): PMW satisfies ε-DP.

Proof: It's an adaptive composition of N AboveThresh and N Laplace Mech, each satisfying ε_0-DP

Theorem (Utility): Choose $N = \frac{16 \log|\mathcal{X}|}{\alpha^2}$, $\varepsilon_0 = \frac{\varepsilon}{2N}$, $\tau = \frac{8 \log\left[\frac{2N+k}{\delta}\right]}{n \varepsilon_0}$

ensures that for any $\alpha \geq \frac{14.53 \left(\log|\mathcal{X}| \cdot \log\frac{3k}{\delta}\right)^{\frac{1}{3}}}{(n\varepsilon)^{\frac{1}{3}}}$

w.p. at least $1-\delta$, ① PMW answers all queries $q_1, \cdots, q_k$.

② $\max_{i \in [k]} |a_i - q_i^T p| \leq \alpha$.

Proof: By Laplace tail bounds and union bound, all Laplace R.V. satisfies $\max_t |w_t| \vee \max_j |v_j| \vee \max_i |z_i| \leq \tau$

There are a total of $2N+k$ Lap R.V.s.

---

for those q_i below threshold

$a_i = q_i^T \tilde{p}_t$, $|a_i - q_i^T p| + v_i \leq \alpha + w_t - 2\tau$

$\Rightarrow |a_i - q_i^T p| \leq \alpha - v_i + w_t - 2\tau \leq \alpha$

---

for those q_i above threshold

$a_i = q_i^T p + z_i \Rightarrow |a_i - q_i^T p| \leq \tau$

By MW regret bound, let's say we answered all K queries at $t=T$.

Assume $T \le N$

$$\sum_{t=1}^{T} \ell_t^T (\hat{p}_t - p) \le 2\sqrt{N \log|\mathcal{X}|}$$

fill in the remaining with $\ell_t = 0$ for all $t \le N$

$$\ell_t^T (\hat{p}_t - p) = \text{sign}\left(q_i^T \hat{p}_t - q_i^T p - z_i \right) \cdot q_i^T (\hat{p}_t - p)$$

$$= \text{sign}\left(q_i^T \hat{p}_t - q_i^T p - z_i\right)\left(q_i^T (\hat{p}_t - p) - z_i + z_i\right)$$

q_i enters MW update $\Leftrightarrow |q_i^T \hat{p}_t - q_i^T p| + \nu_i > \alpha + w_t - 2\tau$

$$\ge \left| q_i^T \hat{p}_t - q_i^T p - z_i \right| - |z_i|$$

$$\ge \left| q_i^T \hat{p}_t - q_i^T p \right| - 2|z_i|$$

$$\ge \alpha + w_t - 2\tau - \nu_i - 2|z_i|$$

$$\ge \alpha - 6\tau \quad \leftarrow (2)$$

(1) and (2) $\Rightarrow$

$$T\alpha - 6T\tau \le 2\sqrt{N \log|\mathcal{X}|}$$

Choose ε_0 such that $6\tau < \frac{\alpha}{2}$

$$\frac{T}{2}\alpha \le T\alpha - 6T\tau \le 2\sqrt{N \log|\mathcal{X}|}$$

$$\Rightarrow T \le \frac{4\sqrt{N \log|\mathcal{X}|}}{\alpha}$$

for $T \le N$, it suffices to choose $N = \frac{16 \log|\mathcal{X}|}{\alpha^2}$

Now let's collect all the constraints and check if we can find ε_0, N such that it works.

- $N = \frac{16 \log|\mathcal{X}|}{\alpha^2}$ ensures $T \le N$
- $\tau \le \alpha$ $\leftarrow$ for Laplace Mech to return a_i with error $\le \alpha$
- $\tau = \frac{8 \log\left(\frac{|\mathcal{X}| + 2N}{\delta}\right)}{\varepsilon_0 n}$
- $2\varepsilon_0 \cdot N = \varepsilon_{\text{budget}}$
- $6\tau \le \frac{\alpha}{2}$ $\leftarrow$ from the above

$$\tau \le \frac{\alpha}{12}$$

So choose $N = \frac{16 \log|\mathcal{X}|}{\alpha^2}$, $\varepsilon_0 = \frac{\varepsilon_{\text{budget}}}{\geq N} = \frac{\varepsilon\alpha^2}{32\log|\mathcal{X}|}$

Now we want to find the smallest α such that this is feasible w.r.t other constraints.

$$\alpha \geq 12\tau = \frac{96 \log\left(\frac{k+2N}{\delta}\right)}{\varepsilon_0 n}$$

$$= \frac{96 \times 32 \log|\mathcal{X}| \log\left(\frac{k+2N}{\delta}\right)}{\alpha^2 \varepsilon \cdot n}$$

$$\alpha^3 \geq \frac{3072 \log|\mathcal{X}| \log\left(\frac{k+2N}{\delta}\right)}{\varepsilon n}$$

it suffices if we choose

$$\alpha^3 = \frac{3072 \log|\mathcal{X}| \log\left(\frac{3k}{\delta}\right)}{\varepsilon n}$$