

Theory of Differential Privacy

Lecture 10:

Private Multiplicative Weights (continued)

EN.601.438/638 Fall 2026

Lydia Zakynthinou

Johns Hopkins University

Oct 5, 2026

Accuracy guarantee of PMW

Thm. Assume $C \cdot \frac{c}{\epsilon n} \log \frac{4kc}{\delta} \leq \frac{\alpha}{2}$, $c = 1 + \left\lceil \frac{4 \log N}{\alpha^2} \right\rceil$ (from NumericSparse)

Then for every $x \in \mathcal{X}^n$ and every adaptive stream of at most k queries, w.p. $1 - \delta$, PMW answers all k queries and

$$\max_{t \in [k]} |f_t(x) - a_t| \leq 3\alpha.$$

Proof strategy

Rely on MW.

- if no update, $b_t = f_t(h^s)$ is accurate.
 - if $\exists$ update, the released v_t is accurate.
 - Each update gives a "valid" MW correction.
 - MW has at most $\frac{4 \log N}{\alpha^2}$ valid corrections.
- $\Rightarrow$ The cap $c = 1 + \left\lceil \frac{4 \log N}{\alpha^2} \right\rceil$ is not exhausted and every query is answered.

} By NumericSparse.

} by MW.

MW Update rule

MWU_η(h, f, v)

Input: distribution h , query $f: X \rightarrow [0, 1]$, approximate answer v , learning rate $\eta \in (0, 1]$

1. Choose cost function $r(a) = \begin{cases} f(a) & , v < f(h) \\ 1 - f(a) & , v \geq f(h) \end{cases}, a \in X$
2. Reweight and normalize:

$$h^{\text{new}}(a) = \frac{h(a) \cdot e^{-\eta r(a)}}{Z} \quad \text{where}$$

$$Z = \sum_{b \in X} h(b) \cdot e^{-\eta r(b)}$$

$f: X \rightarrow [0, 1]$. [e.g.]
 If $f(h) > f(p_x)$
 h puts a lot of mass
 on points $A = \{a: f(a) = 1\}$
 $r(a) = \begin{cases} 1, & \text{if } a \in A \\ 0, & \text{if } a \notin A \end{cases}$

3. Output h^{new} .

Empirical distribution

$$p_x(a) = \frac{1}{n} |\{i: x_i = a\}|$$

for every query $f: X \rightarrow [0, 1]$:

$$\begin{aligned} f(x) &= \frac{1}{n} \sum f(x_i) = \sum_{a \in X} p_x(a) \cdot f(a) \\ f(p_x) &= \sum_{a \in X} p_x(a) \cdot f(a) \end{aligned}$$

Goal:

p_x and h

"close"

PMW

Set $c = 1 + \left\lceil \frac{4 \log N}{\alpha^2} \right\rceil$. Initialize $h^0(a) = \frac{1}{N} \forall a \in X$, Counter $s = 0$.

for each incoming query $f_t, t = 1, \dots, K$:

1. If $s = c$, halt.

2. Set $b_t = f_t(h^s)$

3. Submit query $g_t^+(\cdot) = f_t^+(\cdot) - b_t$ to NumericSparse.

If it returns E , set $v_t = b_t + E$

otherwise submit query $g_t^-(\cdot) = b_t - f_t(\cdot)$ to NumericSparse.

If it returns E , set $v_t = b_t - E$.

4. If both responses are "NO", release $a_t = b_t$ and leave h^s unchanged.

5. If a numeric response occurred, release $a_t = v_t$ and update $h^{s+1} = \text{MWU}_{\alpha/2}(h^s, f_t, v_t)$, $s = s + 1$.

Accuracy Analysis

Lemma 1. (Errors imply cost)

Suppose $|f(h) - f(p_x)| > \alpha$, and $|v - f(p_x)| \leq \alpha/2$.

For the cost function r chosen if MWU:

$$\begin{array}{ccc} r(h) - r(p_x) = |f(h) - f(p_x)| > \alpha \\ \uparrow & & \uparrow \\ \sum_{a \in \mathcal{X}} h(a) \cdot r(a) & & \sum_{a \in \mathcal{X}} p_x(a) \cdot r(a) \end{array}$$

Proof sketch:

- If $f(h) > f(p_x) + \alpha$, then $v \leq \frac{\alpha}{2} + f(p_x) < \alpha + f(p_x) < f(h)$.

So $r = f$. So $r(h) - r(p_x) = f(h) - f(p_x) > \alpha$.

- If $f(p_x) > f(h) + \alpha$, ...

$r = 1 - f$



Potential function:

For a current hypothesis h^s ,
$$\Psi_s = KL(p_x \parallel h^s) = \sum_{\substack{a \in \mathcal{X}: \\ p_x(a) > 0}} p_x(a) \cdot \log \frac{p_x(a)}{h^s(a)}.$$

Fact 1. $\Psi_s \geq 0$

Fact 2. $\Psi_0 = KL(p_x \parallel h^0) = \sum_{a: p_x(a) > 0} p_x(a) \cdot \log(N \cdot p_x(a)) \leq \log N.$

Lemma 2. (one step progress)

for any MNU using cost function $r: \mathcal{X} \rightarrow [0, 1]$ and learning rate $\eta \in (0, 1]$

$$\Psi_s - \Psi_{s+1} \geq \eta (r(h^s) - r(p_x)) - \eta^2.$$

Proof sketch: $Z_s = \sum_{a \in \mathcal{X}} h^s(a) \cdot e^{-\eta r(a)}$, $h^{s+1}(a) = \frac{h^s(a) \cdot e^{-\eta r(a)}}{Z_s}$

$$\begin{aligned} \Psi_s - \Psi_{s+1} &= KL(p_x \parallel h^s) - KL(p_x \parallel h^{s+1}) \\ &= \sum_{a: p_x(a) > 0} p_x(a) \cdot \left(\log \frac{p_x(a)}{h^s(a)} - \log \frac{p_x(a)}{h^{s+1}(a)} \right) = \sum_{a: p_x(a) > 0} p_x(a) \log \frac{h^{s+1}(a)}{h^s(a)} \\ &= \sum_{a: p_x(a) > 0} p_x(a) \cdot \log \frac{e^{-\eta r(a)}}{Z_s} = \sum_{a: p_x(a) > 0} p_x(a) \cdot (-\eta \cdot r(a)) - \sum_{a: p_x(a) > 0} p_x(a) \cdot (\log Z_s) \\ &= -\eta r(p_x) - \log(Z_s) \cdot 1. \end{aligned}$$

Use $e^{-u} \leq 1 - u + u^2$ for $u \in [0, 1]$.

$$\begin{aligned} Z_s &= \sum_a h^s(a) \cdot e^{-\eta r(a)} \leq \sum_a h^s(a) \cdot (1 - \eta r(a) + \eta^2) \\ &= 1 - \eta \sum_a h^s(a) \cdot r(a) + \eta^2 \\ &= 1 - \eta r(h) + \eta^2 \end{aligned}$$

$$\begin{aligned}
\Rightarrow \psi_s - \psi_{s+1} &\geq -\eta r(p_x) - \log z_s \\
&\geq -\eta r(p_x) - \log \left(1 - \overbrace{\eta r(h)}^z + \eta^2 \right) \quad \downarrow \quad \log z \leq z-1 \\
&\geq -\eta r(p_x) - (-\eta r(h) + \eta^2) \\
&= \eta (r(h) - r(p_x)) - \eta^2.
\end{aligned}$$

Theorem 1. h^0 : uniform. Apply PMW with $\eta = \alpha/2$. Suppose every update uses query q satisfying $|f(h^s) - f(p_x)| > \alpha$, $|v - f(p_x)| \leq \alpha/2$. Then the number of updates is

$$m \leq \frac{4 \log N}{\alpha^2} < c.$$

Proof sketch:

$$\text{Lemma 1+2} \Rightarrow r(h^s) - r(p_x) > \alpha \text{ and } \psi_s - \psi_{s+1} > \eta\alpha - \eta^2 = \frac{\alpha^2}{4} \quad (1)$$

$$\text{Sum (1) over } s=0 \dots m-1 : 0 \leq \psi_m \leq \psi_0 - m \frac{\alpha^2}{4} \leq \log N - m \frac{\alpha^2}{4}$$

$$\Leftrightarrow m \geq \frac{4 \log N}{\alpha^2}.$$

Def. Good Event E : all NumericSparse guarantees hold.
 $\Pr[E] \geq 1 - \delta$.

Lemma 3. Condition on E . Every numeric response in PMW satisfies the assumptions of Theorem 1 above.

Proof. If a numeric response occurs then $g(x) \geq \frac{3\alpha}{2} > \alpha$ (by Numeric-sparse guarantees).

For positive test: $g^+(x) = f_t(p_x) - b_t$. So $g^+(x) > \alpha \Leftrightarrow |f_t(p_x) - f_t(h^s)| > \alpha$.

Also, PMW returns $v_t = b_t + E$.

Thus, $|v_t - f_t(p_t)| = |b_t + E - f_t(p_x)| = |E - g^+(x)| \leq \alpha/2$.

So both $|f_t(p_x) - f_t(h^s)| > \alpha$ and $|v_t - f_t(p_x)| \leq \alpha/2$ are satisfied.
Similarly for negative tests. □

Putting the Accuracy guarantee together.

Condition on E . For any round t :

- If both tests return "NO", Numeric Sparse guarantees that for $b_t = f_t(h_t^S)$

$$f_t(p_x) - b_t \leq T + \alpha/2 = 2\alpha + \alpha/2 \leq 3\alpha \quad \text{and} \quad b_t - f_t(p_x) \leq 3\alpha$$

So $|f_t(p_x) - b_t| \leq 3\alpha$ and the response $a_t = b_t$ has $|a_t - f_t(x)| \leq 3\alpha$.

- If a numeric response occurs, PMV returns $a_t = v_t$ and by lemma 3 $|v_t - f_t(x)| \leq \alpha/2 \leq 3\alpha$.
- PMV does not halt. If it did, it would have made c updates, so it would have received c responses. On event E , by lemma 3, all of these queries would have satisfied the conditions of Theorem 1. But Theorem 1 says there can be at most $< c$ updates. Contradiction.

Since w.p. $1-\beta$, event E holds, the Accuracy Guarantee follows.

Remark:

Translating the theorem into sample complexity: $n \geq \frac{\log N}{\epsilon \alpha^3} \log\left(\frac{K \log N}{\beta \alpha^2}\right)$ suffices for accuracy. ▣