On the growth of mistakes in differentially private online learning: A lower bound perspective

Daniil Dmitriev, Kristof Szabo, Amartya Sanyal

ETH AI CENTER · ETH zürich · UNIVERSITY OF COPENHAGEN

Online Game

Let T denote the number of rounds

Adversary selects $x \in \{0,1\}$

and $s = (x_1, \dots, x_T) \in \{x, \perp\}^T$

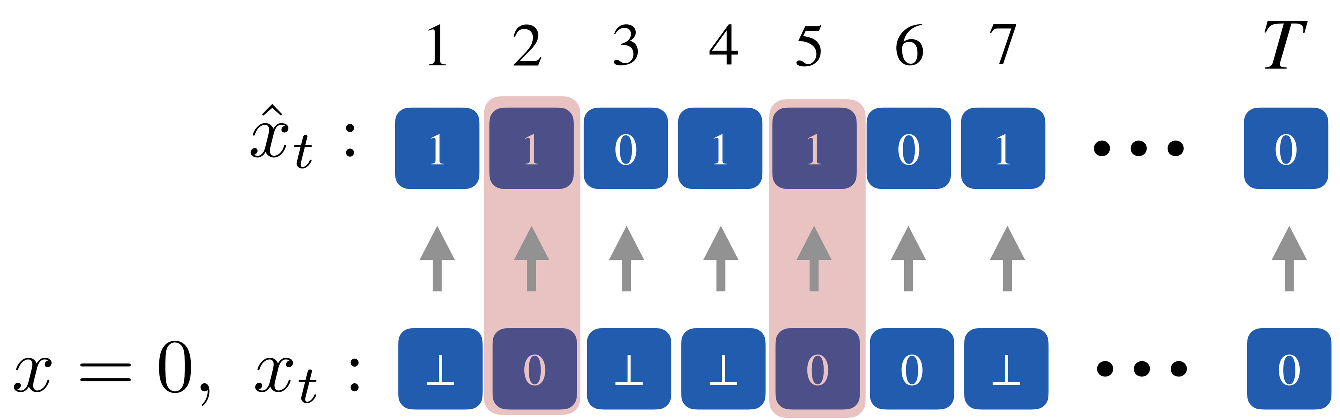
For t in $1 \dots T$:

Learner $\mathcal{A}$ outputs $\hat{x}_t \in \{0,1\}$

Learner $\mathcal{A}$ receives $x_t \in \{x, \perp\}$

Number of mistakes: $M = \sum_{t=1}^{T} \mathbb{I}\{x_t \neq \hat{x}_t \ \& \ x_t \neq \perp\}$

Example:



Differential privacy

Informally:

Learner $\mathcal{A}$ should not strongly depend on any particular x_t

More formally:

For any two inputs s, s' differing at one point: $\mathcal{A}(s) \approx \mathcal{A}(s')$

Most formal (aka definition):

$\mathcal{A}$ is (ε, δ)-private, if $\Pr(\mathcal{A}(s) \in S) \leq e^{\varepsilon} \Pr(\mathcal{A}(s') \in S) + \delta \quad \forall S \subseteq \{0,1\}^T$

To learn **any** function class $\mathcal{H}$ in online setting, the learner must play the game $\mathrm{Ldim}(\mathcal{H}) - 1$ times, where $\mathrm{Ldim}(\mathcal{H})$ is the Littlestone dimension

Main Result

For concentrated or uniform private learners, number of mistakes $\mathbb{E}M = \Omega(\log T)$
(while trivial non-private learner achieves $M \leq 1$)

opposite types of learners

Concentrated:

$$\Pr(\mathcal{A}(\perp, \dots, \perp) = (0, \dots, 0)) = \Omega(1)$$

The only existing upper bound [Golowich and Livni] is concentrated

Uniform:

$$\mathcal{A}(\perp, \dots, \perp)_t \overset{\text{iid}}{\sim} \mathrm{Unif}\{0,1\}$$

Concurrent work [CLNSS'24] shows lower bound against any learner for a particular class

Proof idea

insert point where learner 'expects less'

Assume $\mathcal{A}$ is concentrated and $\varepsilon = \log(3/2)$

Start with $s = (1, \perp, \dots, \perp)$

Let I denote the sequences containing '1' before step $T/2$

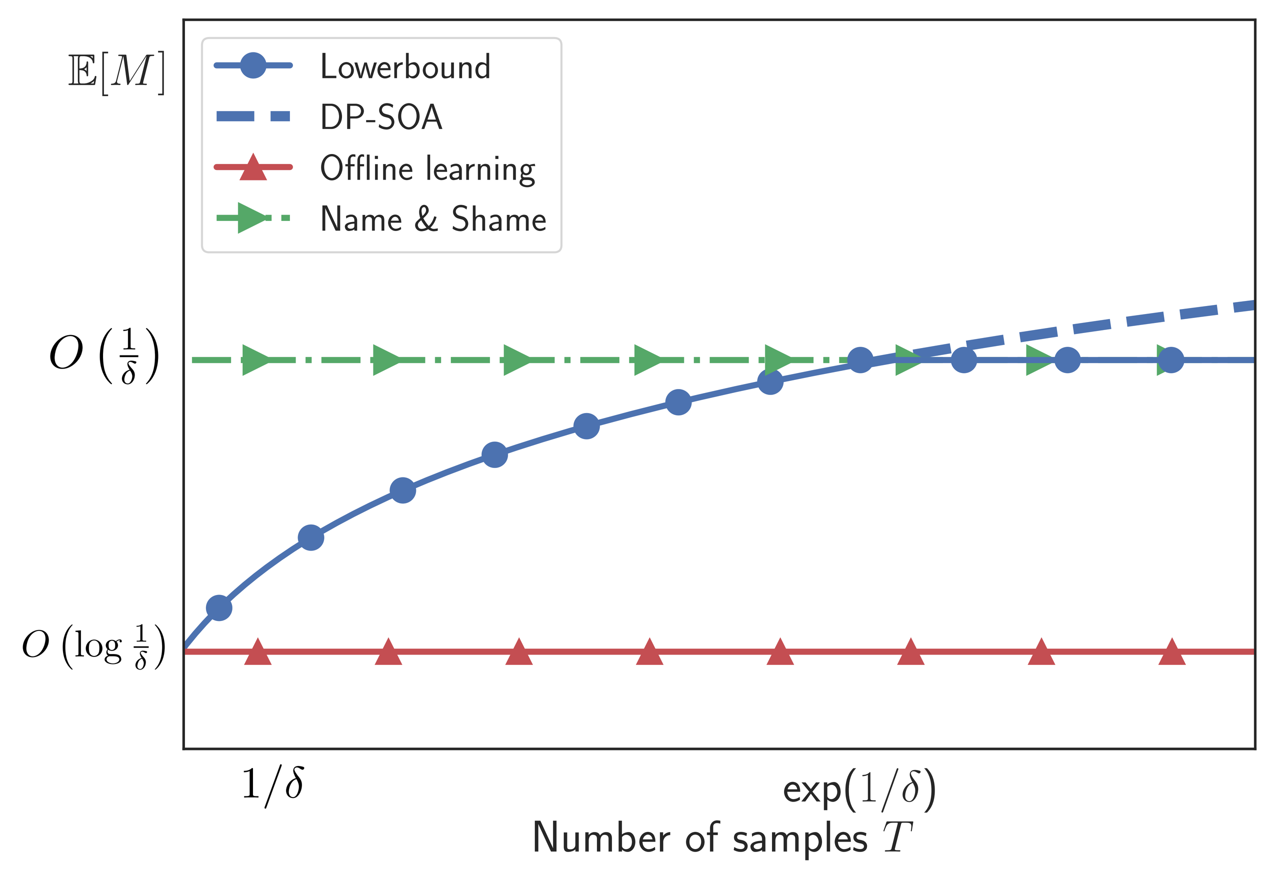
and II denote the sequences containing '1' only after step $T/2$

$\mathcal{A}$ is concentrated $\Longrightarrow \Pr(\mathcal{A}(\perp, \dots, \perp) \in \mathrm{I}) + \Pr(\mathcal{A}(\perp, \dots, \perp) \in \mathrm{II}) =: p \leq \frac{1}{10}$

$\mathcal{A}$ is (ε, δ)-private $\Longrightarrow \Pr(\mathcal{A}(s) \in \mathrm{I}) + \Pr(\mathcal{A}(s) \in \mathrm{II}) \leq \frac{3}{2} p + \delta$

therefore, we recurse to the half I or II, such that $\Pr(\mathcal{A}(s) \in \mathrm{Half}) \leq \frac{3}{4} p + \frac{\delta}{2}$

Overall, we insert $\Omega(\log T)$ '1's, such that $\mathcal{A}$ makes mistake on all of them.



Open Questions

1. Remove any assumptions on $\mathcal{A}$
2. Even for Pure DP ($\delta = 0$) the question is open.

Check it out!

References

[CLNSS'24] Cohen, Edith, Xin Lyu, Jelani Nelson, Tamás Sarlós, and Uri Stemmer. "Lower Bounds for Differential Privacy Under Continual Observation and Online Threshold Queries." The Thirty Seventh Annual Conference on Learning Theory. PMLR, 2024

[GL '21] Golowich, Noah, and Roi Livni. "Littlestone classes are privately online learnable." Advances in Neural Information Processing Systems 34 (2021): 11462-11473.