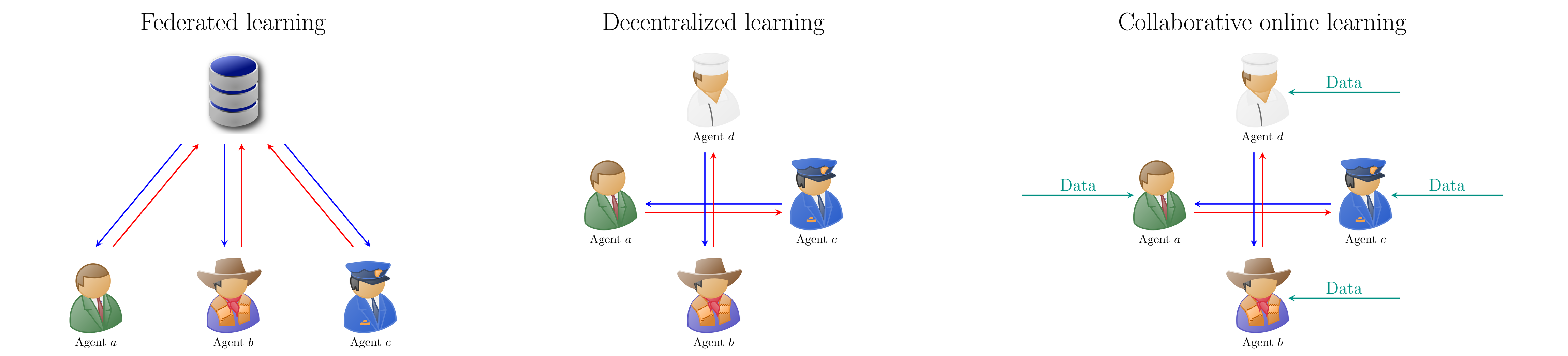
Yauhen Yakimenka[‡], Chung-Wei Weng[†], Hsuan-Yin Lin[†], Eirik Rosnes[†], Jörg Kliewer[‡]
[†]Simula UiB, Norway
[‡]New Jersey Institute of Technology, New Jersey, USA

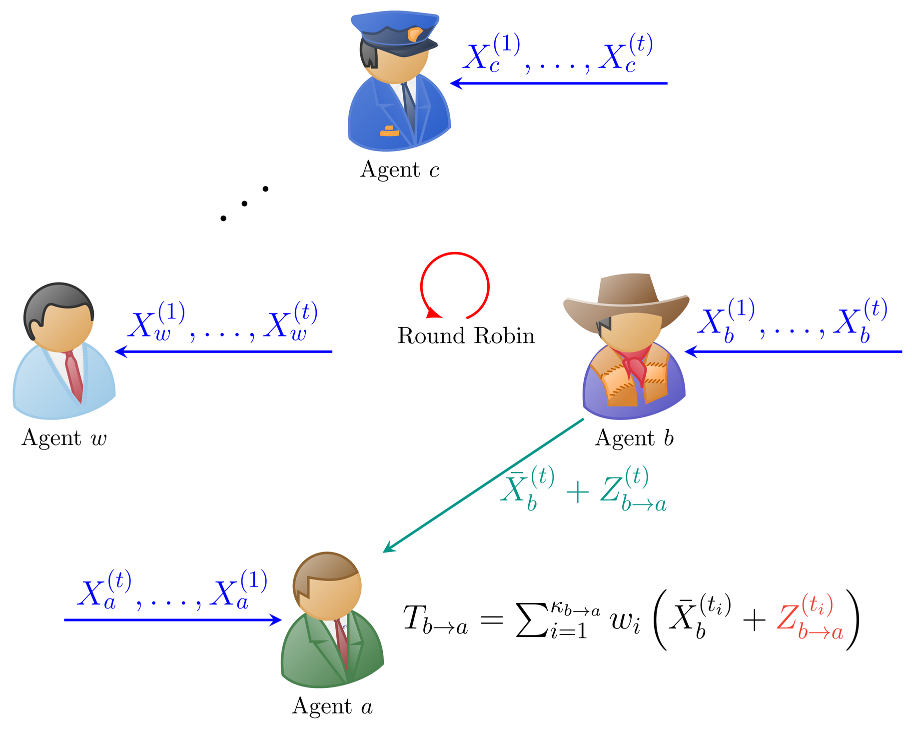
Differentially-Private Collaborative Online Personalized Mean Estimation

1. Motivation



- Collaborative learning has attracted significant attention lately through popular frameworks such as federated learning (FL) [1]
- Focus: Decentralized collaborative online personalized mean estimation [2]
- New: Adding privacy requirement
- Goal: Faster convergence than a fully local approach while providing privacy

2. Problem Statement



- M independent agents
- Each agent a wants to estimate the mean of its sample $X_a^{(1)}, X_a^{(2)}, \ldots \in \mathcal{X}_a \subset \mathbb{R}$
- $X_a^{(i)} \sim \mathcal{D}_a$ with bounded support $\mathcal{X}_a$ with an (unknown) mean μ_a and known/unknown standard deviation $\sigma_a < \infty$
- For some users a and b, $\mu_a = \mu_b$, and a and b belong to the same *class*
- At each time step t, agent a receives $X_a^{(t)}$, updates its sample mean $\bar{X}_a^{(t)}$, and also chooses another agent b to query
- Agent b then sends its current sample mean to a, but privatized: $\bar{X}_b^{(t)} + Z_{b\to a}^{(t)}$
- A particular construction of the noise $Z_{b\to a}^{(t)}$ is a *private release mechanism*
- Agent a computes $T_{b\to a} = \sum_{i=1}^{\kappa_{b\to a}} w_i \left(\bar{X}_b^{(t_i)} + Z_{b\to a}^{(t_i)}\right)$ (estimate of agent b's mean)
- Decision rule of agent a: Agent b has the same distribution mean as me ($\chi_a^{(t)}(b;\theta_t) = 1$) if $\left|\bar{X}_a^{(t)} - T_{b\to a}\right| < \Phi_{t,\nu}^{-1}\left(1 - \frac{\theta_t}{2}\right)\sqrt{\frac{\widehat{\sigma_a^2}}{t} + \widehat{\mathsf{Var}}[T_{b\to a}]}$ (hypothesis testing; student t-distribution)

4. Contributions [3]

- Two (online) differential privacy (DP) mechanisms inspired by the ones in [4] are proposed
- A theoretical convergence analysis showing convergence
- The best scheme performs comparably to ideal performance where all data is public
- Compared to [3]: σ_a is assumed unknown and estimated for all agents a
  - Var-Est-1: A privatized partial sample variance is released
  - Var-Est-2: Variance is estimated from the already released privatized sample means

5. Differential Privacy Mechanisms

Privatized version of $\bar{X}_b^{(t)}$ using so-called *p-sums*:

$$\bar{X}_b^{(t)} + Z_{b\to a}^{(t)} = \frac{X_b^{(1)} + \cdots + X_b^{(t)}}{t} + Z_{b\to a}^{(t)}$$
$$= \frac{\sum_{i=1}^{\tau_1} X_b^{(i)} + Z_{b\to a}^{(1:\tau_1)} + \sum_{i=\tau_1+1}^{\tau_2} X_b^{(i)} + Z_{b\to a}^{(\tau_1+1:\tau_2)} + \cdots + \sum_{i=\tau_{\kappa-1}+1}^{t} X_b^{(i)} + Z_{b\to a}^{(\tau_{\kappa-1}+1:t)}}{t}$$

- PM-I: Split $[1:t_\kappa]$ into $[1:t_1], [t_1+1:t_2], \ldots, [t_{\kappa-1}+1:t_\kappa]$
- PM-II: Join the subsums of PM-I into larger subsums according to the binary representation of κ

3. Our Approach

Algorithm 1: Private-ColME

Input: agent a

Output: $\mu_a^{(t_{\max})}$

$\forall b \in [M] \setminus \{a\} : T_{b\to a} \leftarrow 0, \kappa_{b\to a} \leftarrow 0$

$\mathcal{C}_a^{(0)} \leftarrow [M]$

for $t = 1, 2, \ldots, t_{\max}$ **do**

  `// Receive`

  Receive sample $X_a^{(t)} \sim \mathcal{D}_a$

  $\bar{X}_a^{(t)} \leftarrow \bar{X}_a^{(t-1)} \times \frac{t-1}{t} + X_a^{(t)} \times \frac{1}{t}$

  `// Query`

  $b \leftarrow \texttt{choose\_agent}\left(\mathcal{C}_a^{(t-1)}, [M]\right)$

  $\kappa_{b\to a} \leftarrow \kappa_{b\to a} + 1$

  $T_{b\to a} \leftarrow \sum_{i=1}^{\kappa_{b\to a}} w_i \left(\bar{X}_b^{(t_i)} + Z_{b\to a}^{(t_i)}\right)$

  Update $\widehat{\sigma_a^2}$, $\widehat{\sigma_b^2}$, and $\widehat{\mathsf{Var}}[T_{b\to a}]$

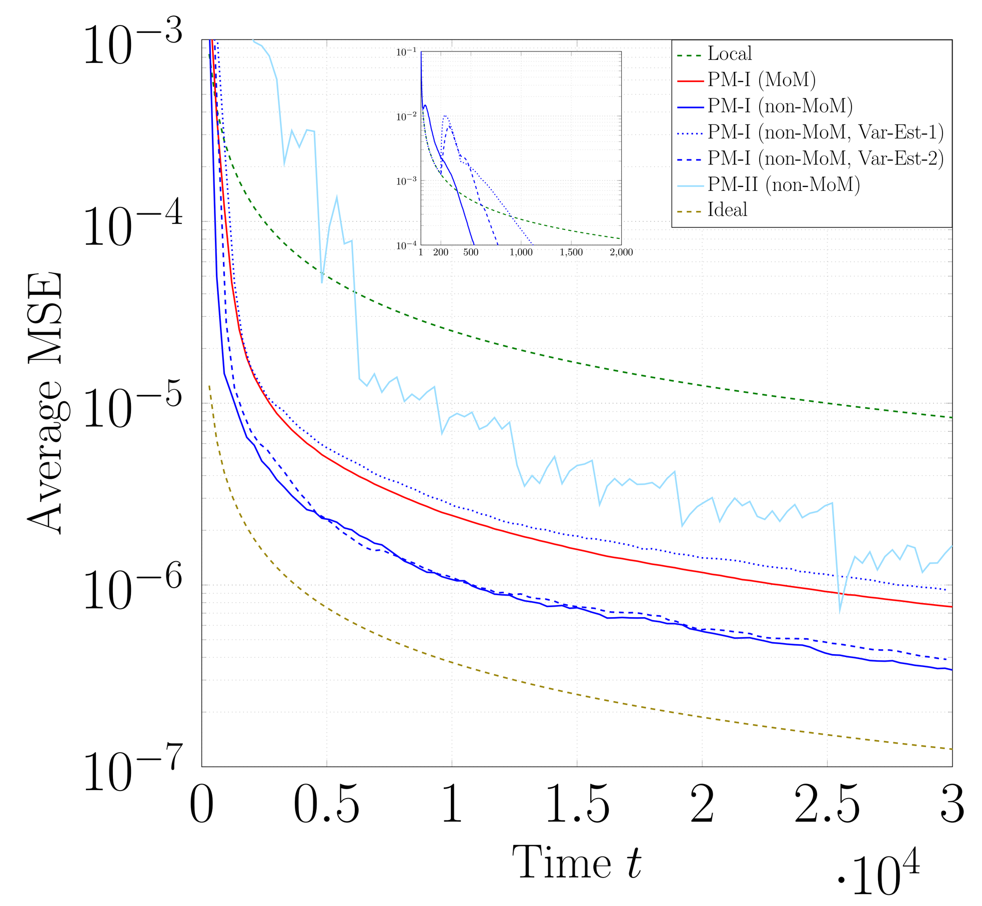
  `// Estimate`

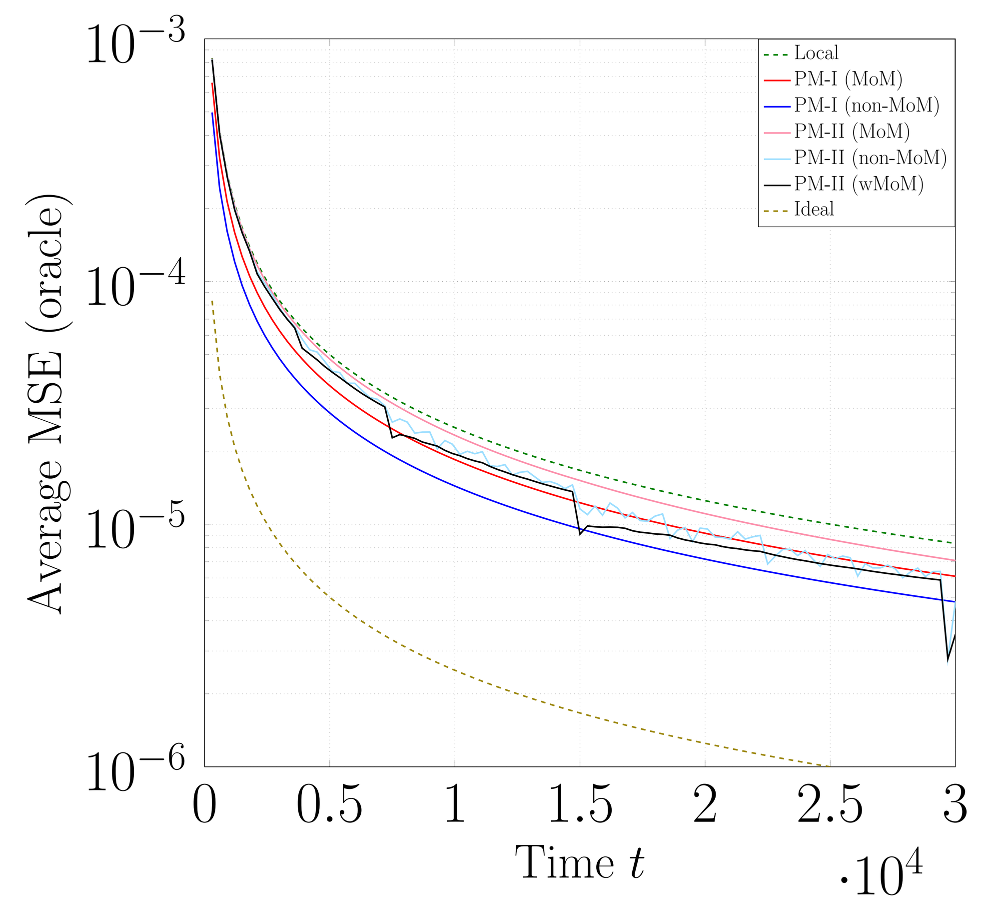
  $\mathcal{C}_a^{(t)} \leftarrow \{b \in [M] : \chi_a^{(t)}(b;\theta_t) = 1\}$

  $\mu_a^{(t)} \leftarrow \alpha_{a\to a}^{(t)} \bar{X}_a^{(t)} + \sum_{b \in \mathcal{C}_a^{(t)} \setminus \{a\}} \alpha_{b\to a}^{(t)} T_{b\to a}$

return $\mu_a^{(t_{\max})}$

6. Results





- 200 and 30 agents, resp., and three classes
- Uniform data with class-dependent means
- DP: $\epsilon = 1$ with $\delta = 10^{-6}$ (Gaussian mechanism)
- Decision rule: $\theta_t = {}^{0.05}/_{\ln(t+1)}$

References

[1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artif. Intell. Statist. (AISTATS)*, Ft. Lauderdale, FL, USA, Apr. 20–22, 2017, pp. 1273–1282.

[2] M. Asadi, A. Bellet, O.-A. Maillard, and M. Tommasi, "Collaborative algorithms for online personalized mean estimation," *Trans. Mach. Learn. Res.*, 2022.

[3] Y. Yakimenka, C.-W. Weng, H.-Y. Lin, E. Rosnes, and J. Kliewer, "Differentially-private collaborative online personalized mean estimation," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Tapei, Taiwan, Jun. 2023.

[4] T.-H. H. Chan, E. Shi, and D. Song, "Private and continual release of statistics," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 3, pp. 1–24, Nov. 2011.