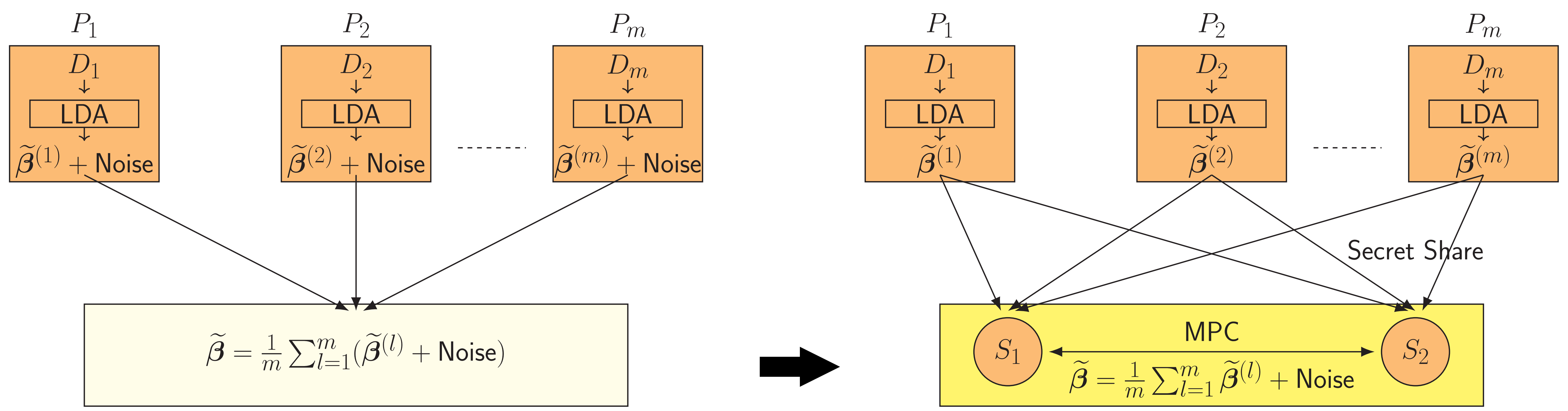
Aggregating Private Sparse Learning Models Using Multi-Party Computation

Lu Tian*, Bargav Jayaraman*, Quanquan Gu, David Evans

University of Virginia

Secure Aggregation with Differential Privacy

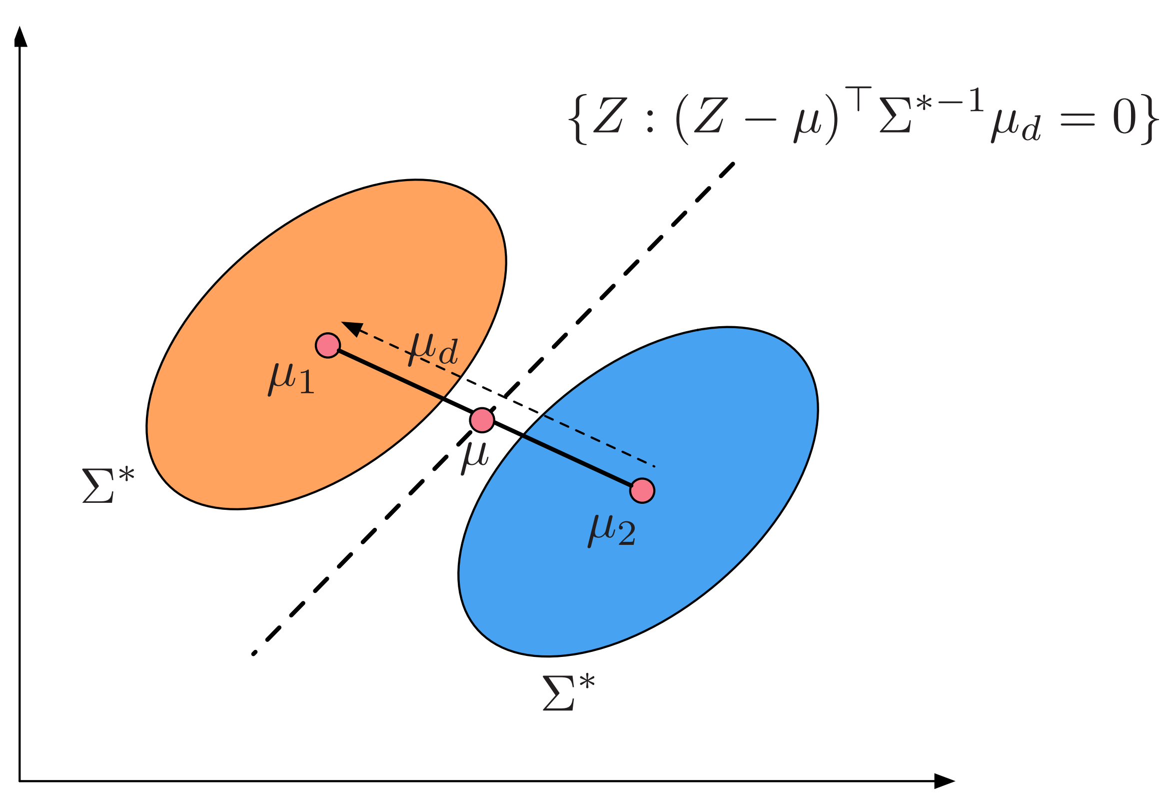
We consider the problem of privately learning a sparse model across multiple sensitive datasets. Individual models are locally learned and privately aggregated using secure multi-party computation (MPC).



(a) Aggregation of noisy models **(b) Adding noise in MPC**

Adding privacy-preserving noise after aggregation, instead of before, leads to more accurate models.

Distributed Sparse Learning



We focus on distributively estimating

$$\boldsymbol{\beta}^* := \boldsymbol{\Sigma}^{*-1}\boldsymbol{\mu}_d$$

Each party estimates biased discriminant function

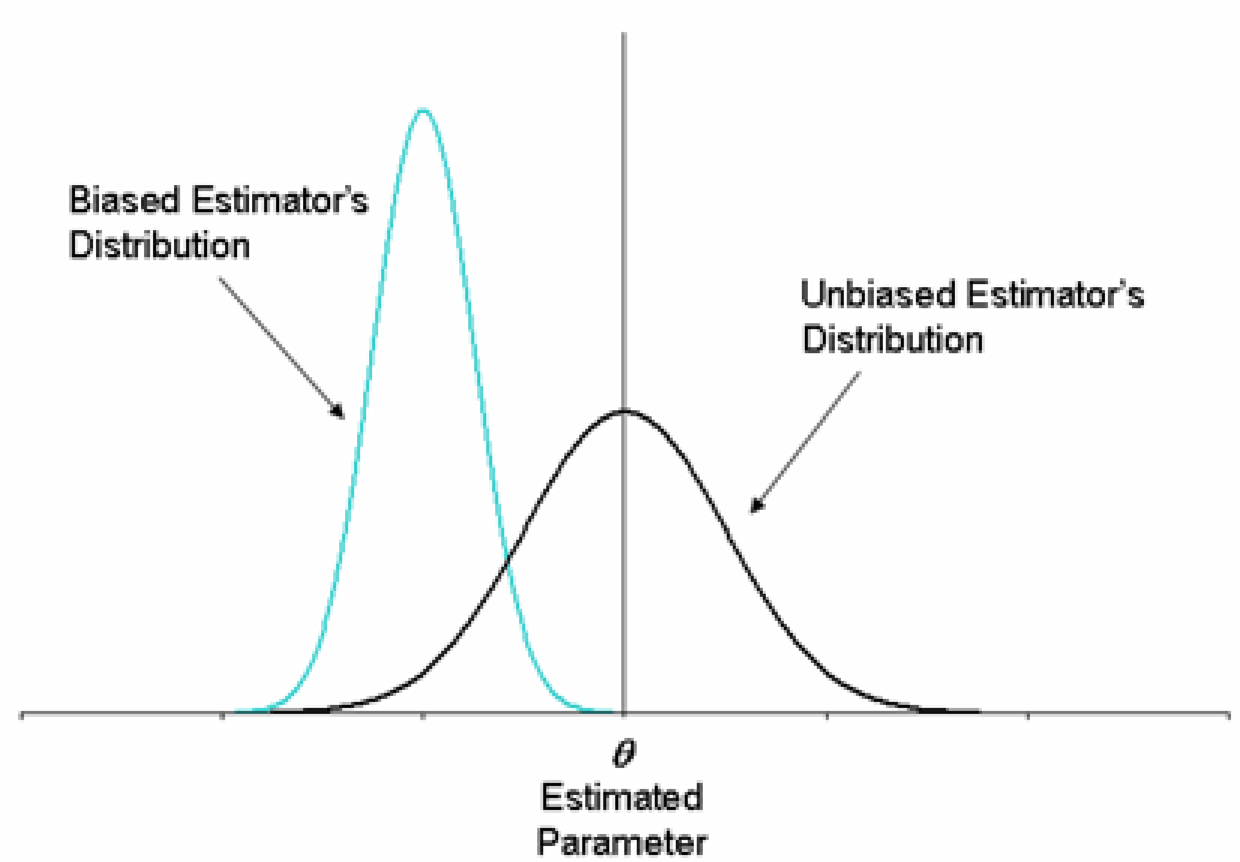
$$\widehat{\boldsymbol{\beta}}^{(l)} = \underset{\boldsymbol{\beta}}{\text{argmin}} \|\boldsymbol{\beta}\|_1$$

$$\text{subject to } \|\widehat{\boldsymbol{\Sigma}}^{(l)}\boldsymbol{\beta} - \widehat{\boldsymbol{\mu}}_d^{(l)}\|_\infty \le \lambda$$

Debiasing Sparse Models

Each party estimates unbiased discriminant function

$$\widetilde{\boldsymbol{\beta}}^{(l)} = \widehat{\boldsymbol{\beta}}^{(l)} - \widehat{\boldsymbol{\Theta}}^{(l)\top}(\widehat{\boldsymbol{\Sigma}}^{(l)}\widehat{\boldsymbol{\beta}}^{(l)} - \widehat{\boldsymbol{\mu}}_d^{(l)})$$



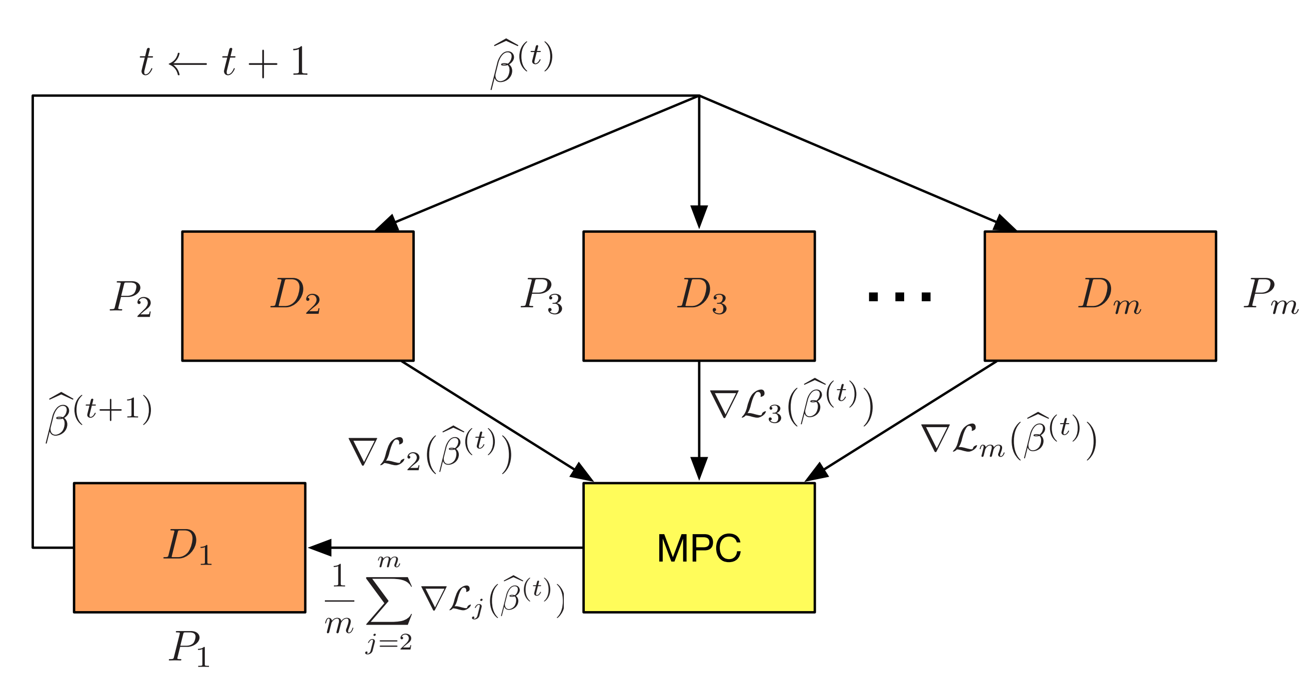
Experiment Setting

Synthetic Dataset: Number of parties varies from 20 to 100. The dimensionality of data is set as 200, with each party having 200 data instances generated from two Gaussian distributions.
Hospital Dataset: 920 patient records from 4 hospitals. Records contain personal information like age and gender, and clinical information such as laboratory test results. We aim at predicting whether a patient has heart disease.

Experiments

Dataset	m	Misclassification Rate		
		Centralized LDA	Naive Averaged	Our Approach
Synthetic	20	0.168 ± 0.002	0.240 ± 0.003	0.182 ± 0.003
Synthetic	60	0.166 ± 0.001	0.240 ± 0.002	0.179 ± 0.002
Synthetic	100	0.165 ± 0.001	0.240 ± 0.002	0.179 ± 0.001
Hospital	4	0.208 ± 0.012	0.329 ± 0.035	0.220 ± 0.017

Future Work: Iterative Learning



$\nabla L_j(\cdot)$'s are gradient of loss functions.

*Both the authors contributed equally