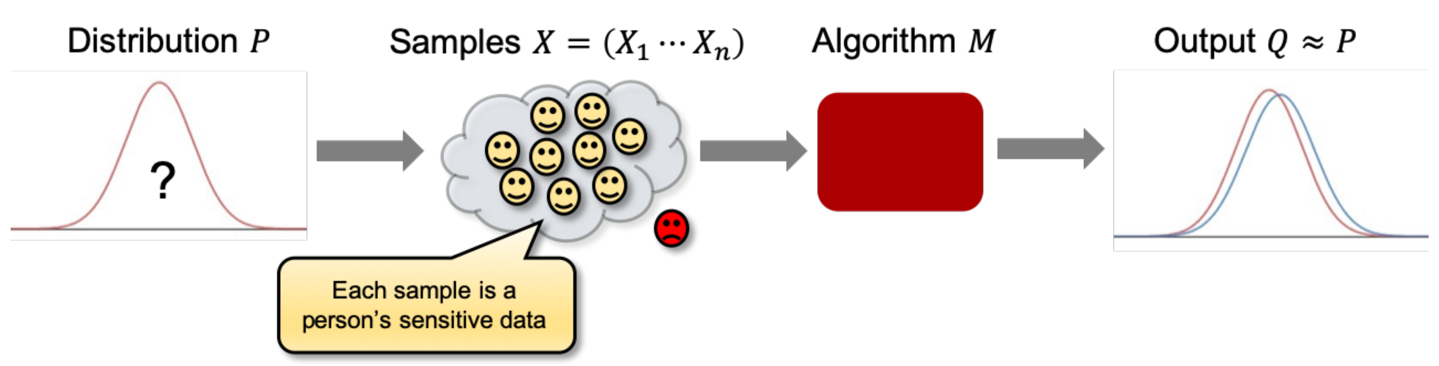
Differentially Private Algorithms for Learning Mixtures of Separated Gaussians

Gautam Kamath[1], Or Sheffet[2], Vikrant Singhal[3], and Jonathan Ullman[3]

[1] University of Waterloo; [2] Bar-Ilan University; [3] Northeastern University

1 - Private Distribution Learning



- **Goal:** "Learn" P while "hiding" the sample X "for free": (1) sample efficiency, (2) time efficiency, and (3) minimal constraints on distribution parameters
- **Our Work:** Learning mixtures of high-dimensional Gaussians with Differential Privacy
  - New private annulus finding algorithm (technical strengthening of [NS'18])
  - New guarantees for private PCA
  - New private Gaussian clustering algorithm
  - Beats Subsample & Aggregate ((1) only works for spherical Gaussians, (2) has high sample complexity)

2 - Learning Gaussian Mixtures

α-Learning: Given a mixture of k Gaussians $\{G_i \equiv N(\mu_i, \Sigma_i)\}_{i=1}^k$ in $\mathbb{R}^d$ with mixing weights $\{w_i\}_{i=1}^k$, $\forall i$, estimate G_i to within α in TV distance and w_i to within $O\left(\frac{\alpha}{k}\right)$.

Parameter Constraints: $\forall i$, $\|\mu_i\|_2 \leq R$, $\mathbb{I} \preccurlyeq \Sigma_i \preccurlyeq K\mathbb{I}$, and $w_i \geq w_{min}$

Separation Condition: $\forall i, j$,

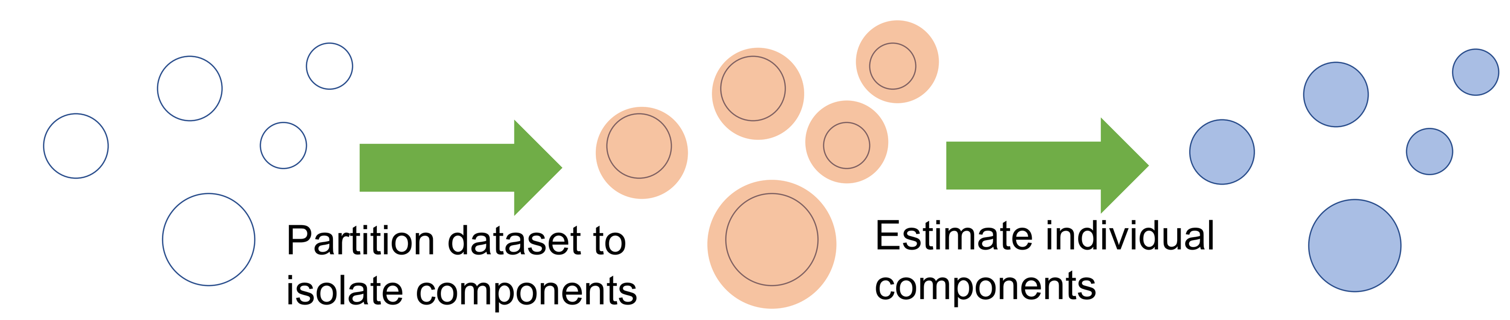
$$\|\mu_i - \mu_j\|_2 \gtrsim \left(\|\Sigma_i\|_2 + \|\Sigma_j\|_2\right)\left(\sqrt{k} + \frac{1}{\sqrt{w_i}} + \frac{1}{\sqrt{w_j}}\right).$$

Theorem: $\exists$ (ε, δ)-DP alg for α-learning mixtures of Gaussians that has sample complexity:

$$\tilde{O}\left(\frac{d^2}{\alpha^2 w_{min}} + \frac{d^2}{\alpha w_{min} \varepsilon} + \frac{k^{9.06} d^{1.5}}{w_{min} \varepsilon}\right).$$

3 - Case: Beginner

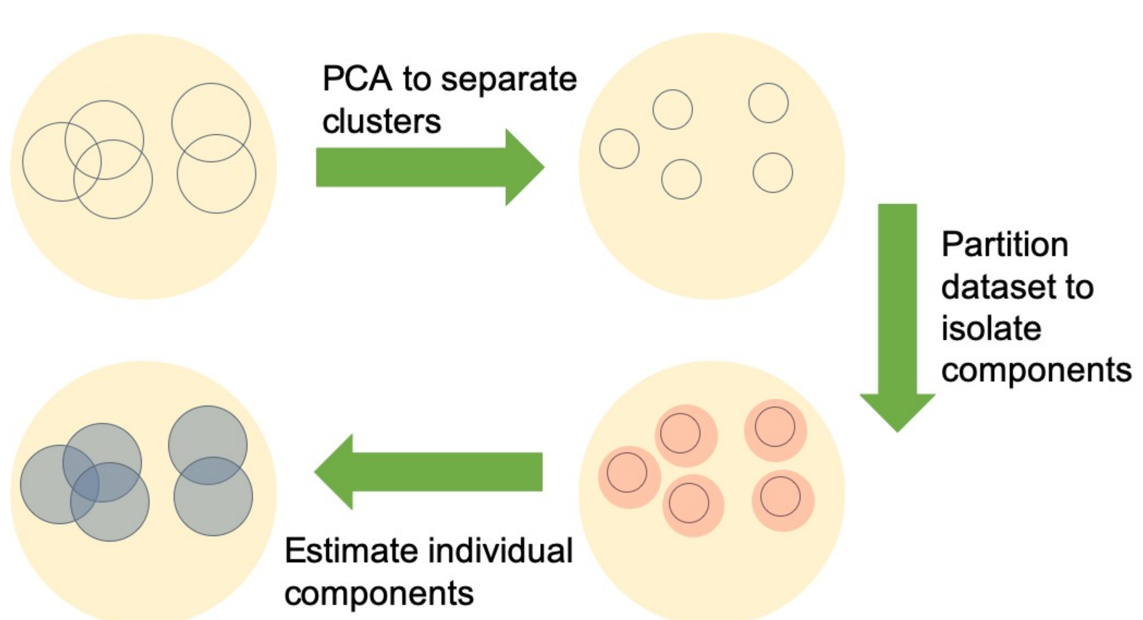
- Means separated by $\mathbf{\Omega}(\sqrt{\boldsymbol{d}})$ (clusters far from each other)
- Uniform mixing weights



- **Step 1:** Private clustering algorithm from [NS'18]
- **Step 2:** Private Gaussian learner from [KLSU'19]

4 - Case: Intermediate

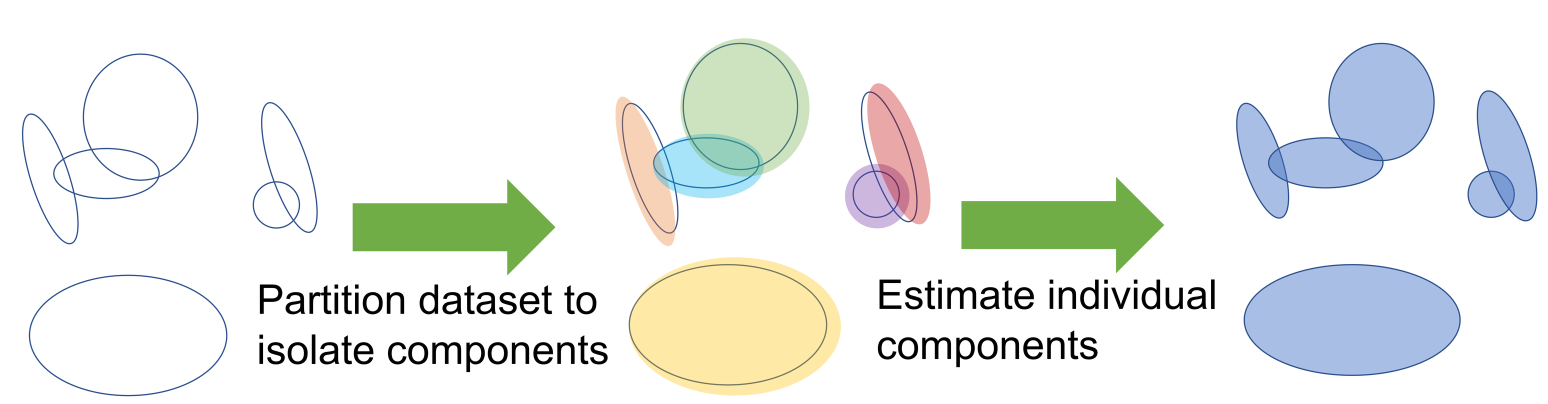
- Means separated by $\mathbf{\Omega}(\sqrt{\boldsymbol{k}})$
- Spherical Gaussians: variances within $\Theta(1)$ of each other
- Means lie in a ball of radius $O(k\sqrt{d})$ around origin
- Uniform mixing weights



- **Step 1:** Private PCA
  - Shrinks Gaussians whilst maintaining separation
- **Step 2:** Private clustering algorithm from [NS'18]
- **Step 3:** New Private Spherical Gaussian learner

5 - Case: Pro

- Mixture satisfies all conditions in Panel 2

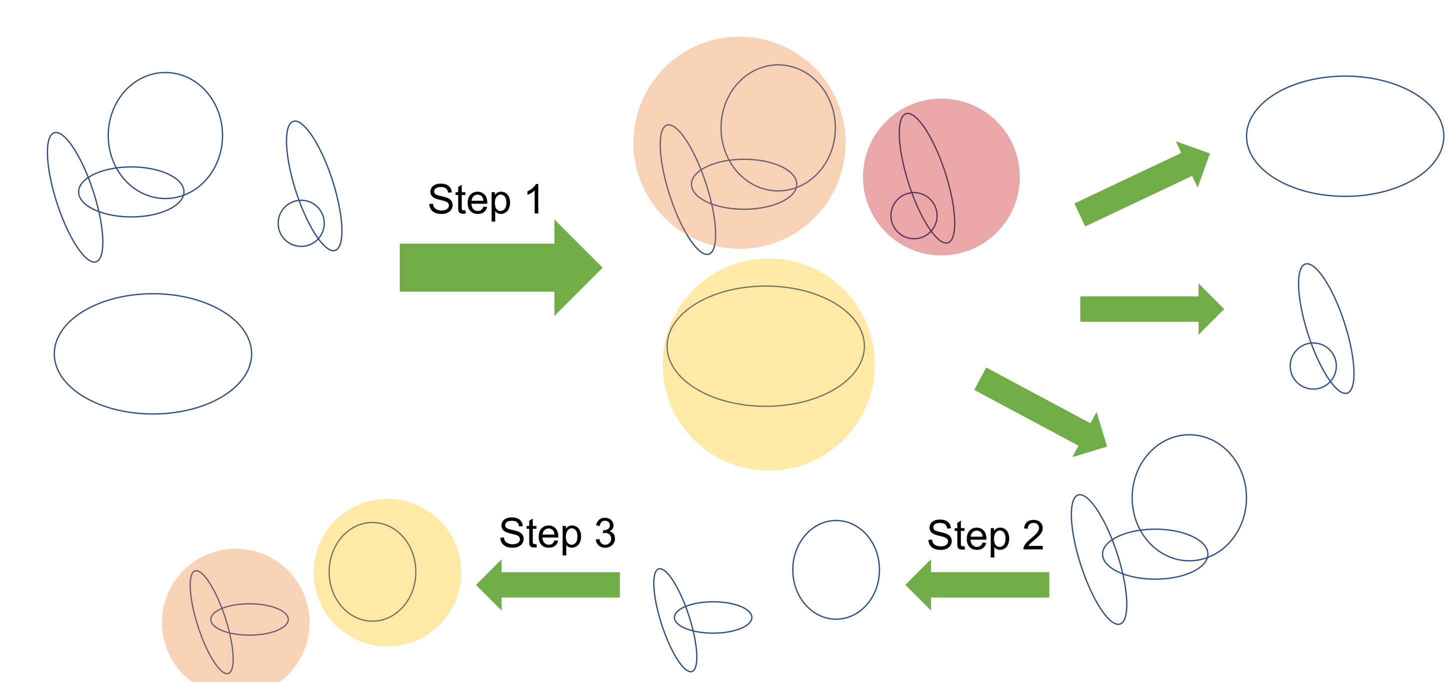


- **Step 1:** Recursive Private Partitioner (clustering)
- **Step 2:** Adaptation of Gaussian learner from [KLSU'19] for when few points could be lost in Step 1

Recursive Private Partitioner (Key Ideas):

- Every group of nearby clusters could be treated as independent sub-problem
- Want to isolate such groups in small balls to reduce sensitivity for later
- Largest cluster in each group can be separated at low cost

6 - Case: Pro (Clustering)



- **Step 1:** Isolate distant groups of clusters within disjoint balls of radius $O(k\sqrt{d})$ using private annulus finding alg
- **Steps 2:** Separate large Gaussians from smaller ones using private PCA
- **Steps 3:** Isolate largest Gaussian from the remaining ones using algorithm in Step 1
- **Recurse on the sub-problems**