

Revisiting EM-Based Locally Differentially Private Protocols

基于期望最大化的本地差分隐私协议分析

Published in Network and Distributed System Security Symposium (NDSS 2025). (CCF A)

叶宇桐, 王天豪, 张敏, 冯登国

联系人: 叶宇桐 yutong2017@iscas.ac.cn

研究背景

本地差分隐私 (LDP) 作为分布式隐私计算的核心范式, 如图1所示, 允许用户直接在终端设备上向敏感数据 x 注入可控噪声, 再将扰动后的数据 \tilde{x} 上传至不可信服务器, 面向不同数据类型进行相应任务的分析。然而, LDP 的强隐私保障以数据效用损耗为代价, 即隐私级别越高, 噪声越多, 在噪声数据集上的分析失真越多。现有聚合分析方法大多追求聚合结果为真实目标的无偏估计 (如采用矩阵逆变换), 虽理论无偏但因噪声过强常产生负值或总和偏差 (如图2 (a) 所示)。聚合分析方法的优化和选取, 仍然需要深入探讨。

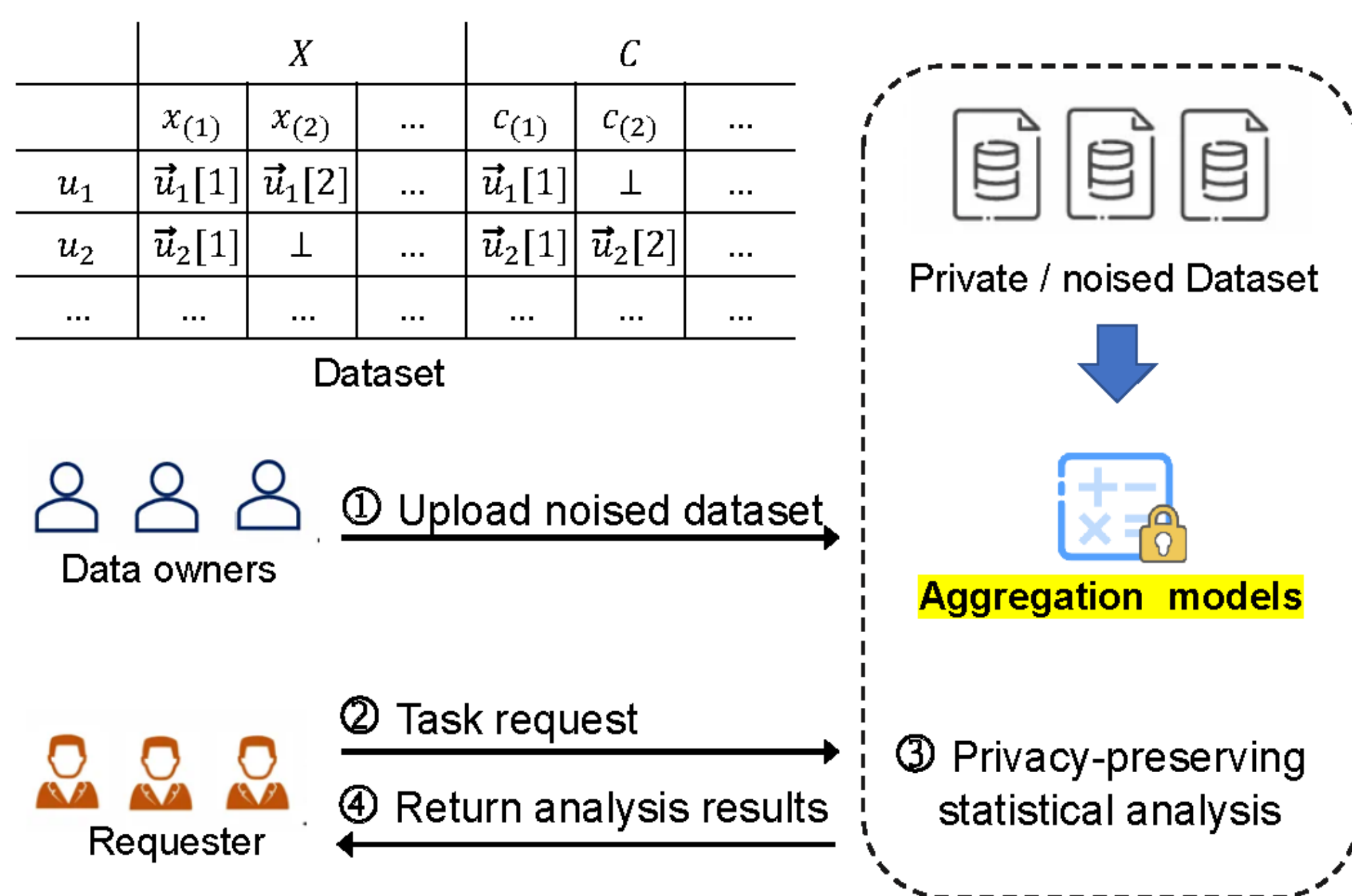


图1. LDP协议框架

EM-based 方法

另一类聚合分析方法是寻找最可能生成观测扰动数据 $\{\tilde{x} \dots\}$ 的假设分布 (如图2 (b) 所示), 再用分布获取目标估计结果。生成假设分布的过程采用期望最大化算法 (EM), 但本文发现基于 EM 的方式易过拟合于扰动噪声, 并未推广到多种任务。为此, 本文提出一种引入高斯混合模型的归约思想 (Mixture Reduction, MR), 提出动态分量合并框架来处理噪声数据, 即在EM迭代中持续剔除低权重分量, 降低模型复杂度并加速收敛。

Take frequency estimation task for example, unbiased method is to derive function for each item,

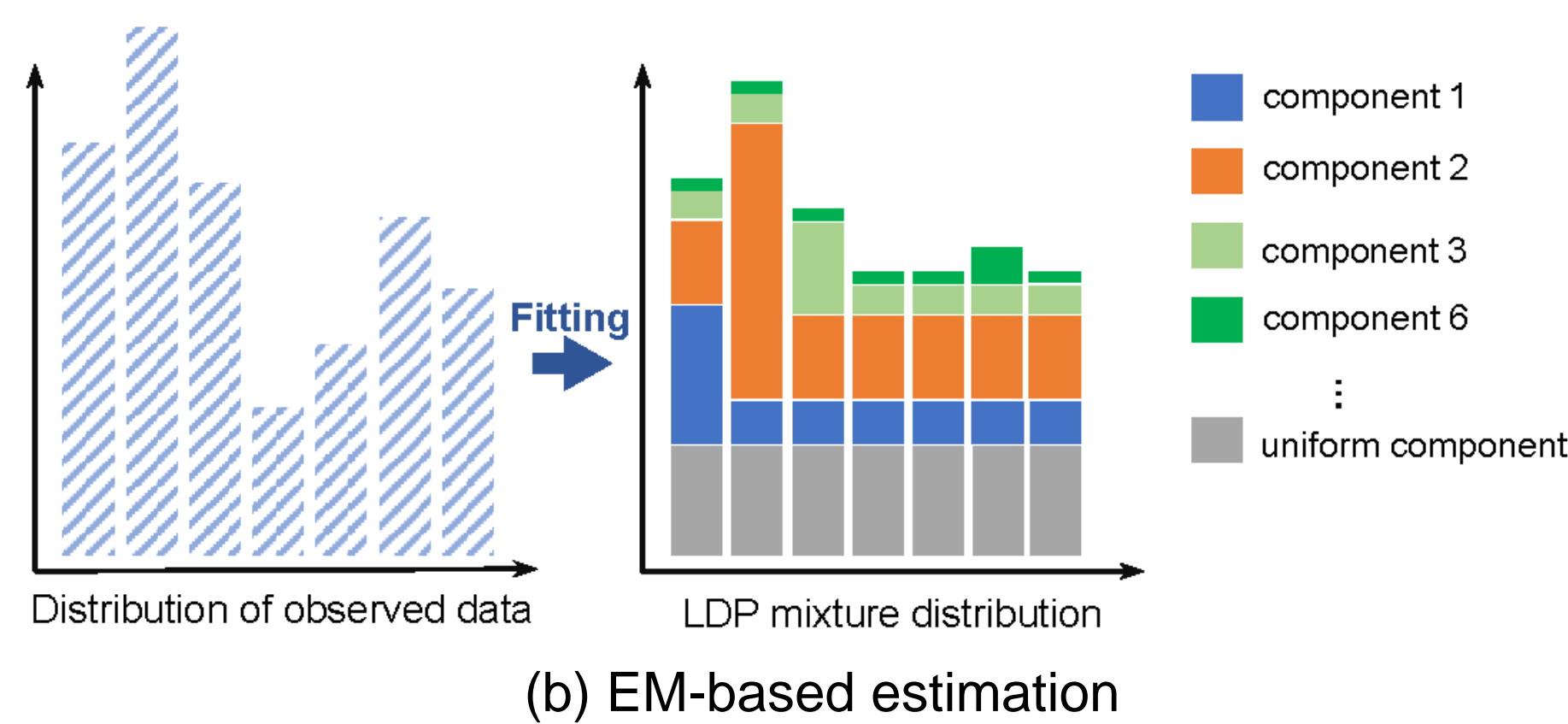
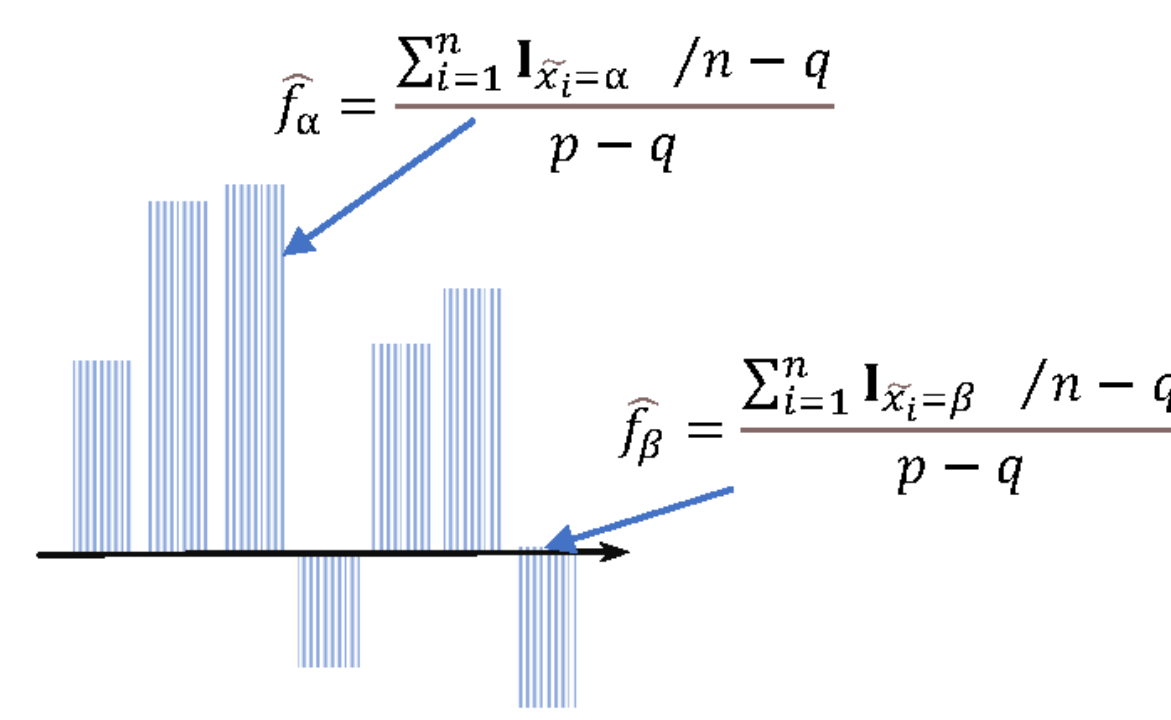


图2. 聚合分析方法对比

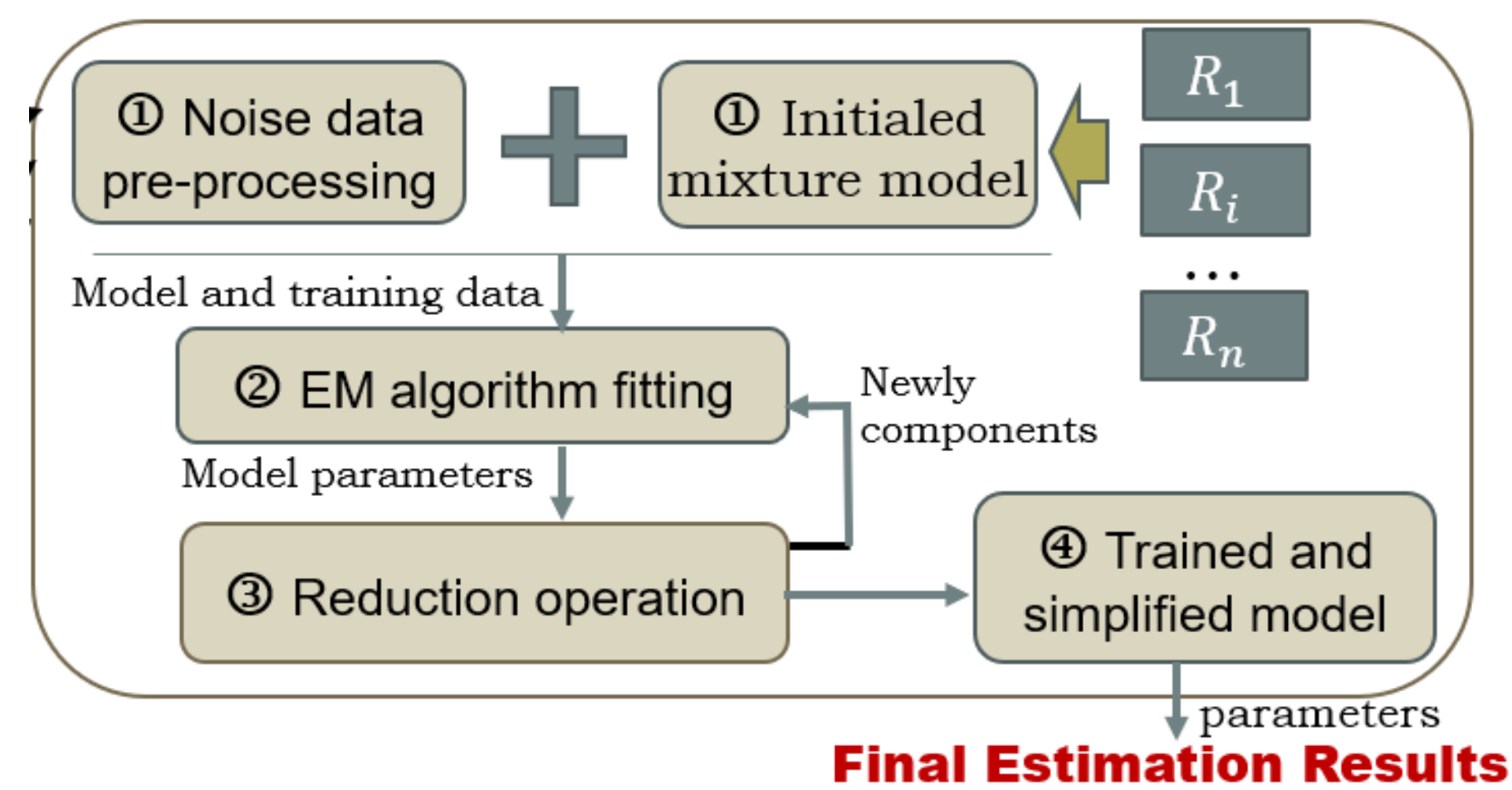


图3. MR方法的工作流程

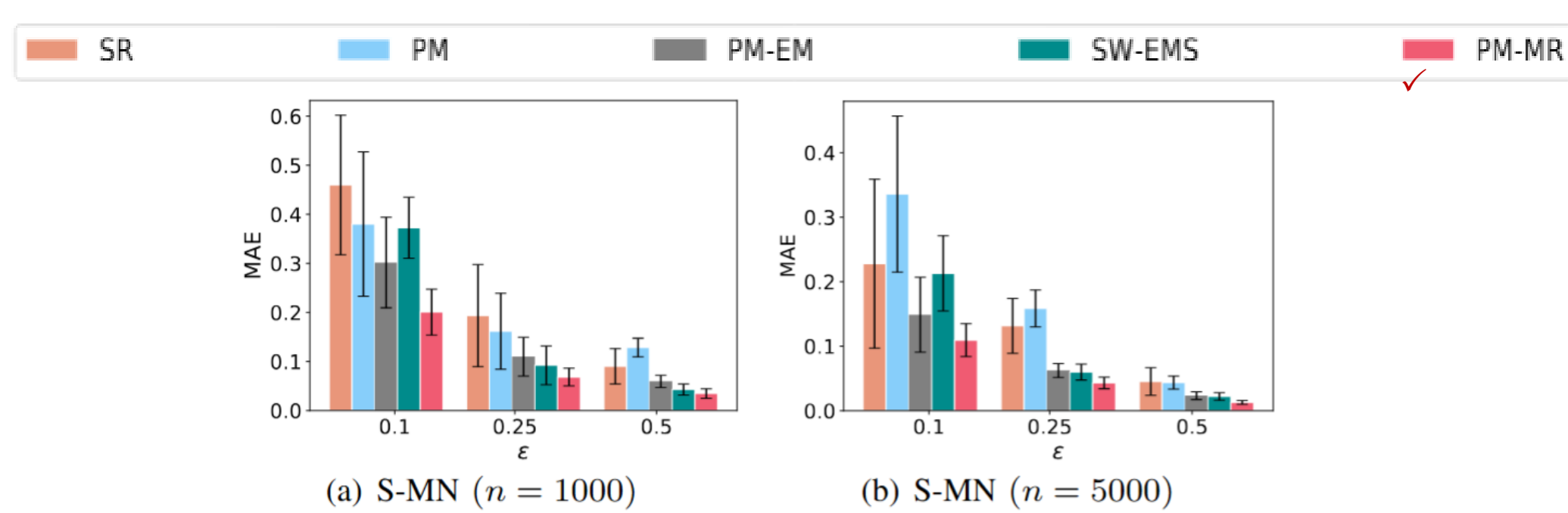


图4. MR在均值统计任务与SOTA对比

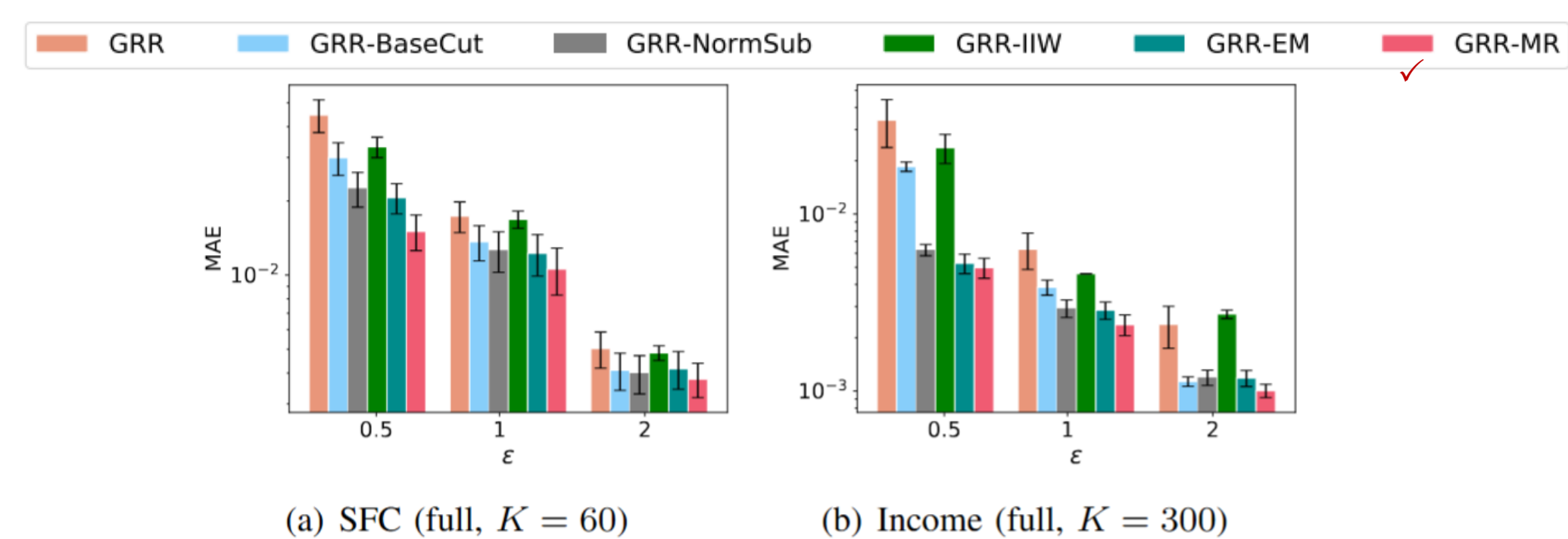


图5. MR在频率估计任务与SOTA对比

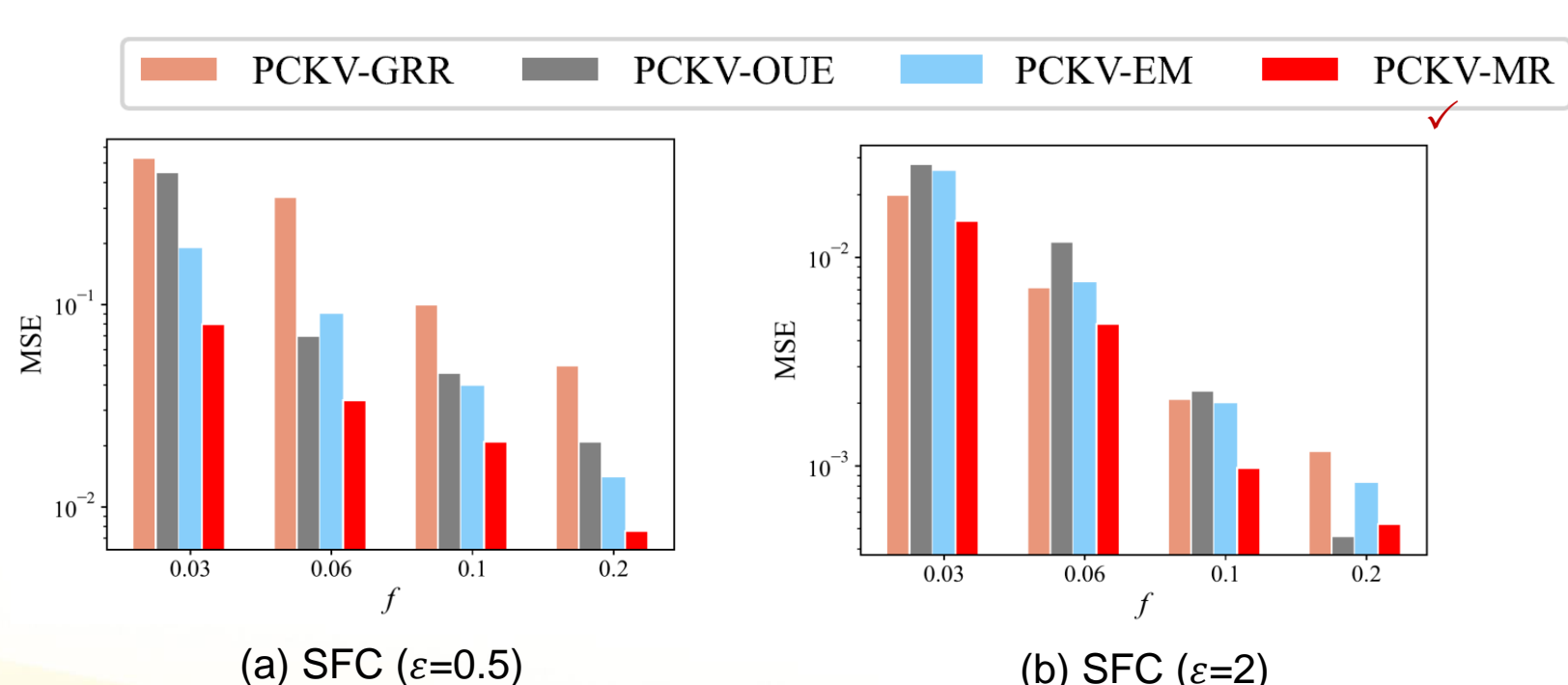


图6. MR在条件分析任务与SOTA对比

研究方法

定义LDP混合模型拟合噪声数据, 通过对噪声性质的分析为模型拟合期间加入规约简化步骤:

- 泛化过程(图3 ①), 首先建立 LDP 混合模型

$$\phi(\tilde{x}; w, \alpha) = \sum_k w_k \Pr[\psi_\epsilon(\alpha_k) = \tilde{x}]$$

- $\Pr[\psi_\epsilon(\alpha_k) = \tilde{x}]$: PMF, 或是扰动噪声的函数

目标: $\arg \max_{\hat{w}} \mathcal{L}(\hat{w})$ s.t. $\sum \hat{w}_k = 1, \hat{w}_k \geq 0$

E-Step:

$$Y_{ik} \leftarrow \frac{\hat{w}_k \Pr[\Psi_\epsilon(\alpha_k) = \tilde{x}_i]}{\sum_{j=1}^K \hat{w}_j \Pr[\Psi_\epsilon(\alpha_j) = \tilde{x}_i]}$$

M-Step:

$$\hat{w}_k \leftarrow \frac{1}{n} \sum Y_{ik}$$

- 执行规约简化步骤(图3 ③④), 混合成分分布的合并 $(w_{12}, \Psi_\epsilon(\alpha_{12})) \leftarrow \{(w_1, \Psi_\epsilon(\alpha_1)), (w_2, \Psi_\epsilon(\alpha_2))\}$ 。迭代

终止条件: $\text{BIC} = -2 \log(\mathcal{L}) + K' \log(n)$

实验结果

本文在多个基准数据集和三大主流分析任务上进行了广泛的评估。实验结果如图4~6所示, 本文提出的MR方案在MSE、MAE、Wasserstein distance等指标上, MR在隐私级别高的情况下都始终优于现有EM和无偏估计, 在隐私级别低时几乎等价于无偏估计。