Correlated Privacy Mechanisms for Differentially Private Distributed Mean Estimation

Distributed mean estimation (DME) is a fundamental building block in a number of applications such as federated learning, and distributed stochastic gradient descent. Differentially private DME (DP-DME) refers to the setting where a central server aims to estimate the mean of d dimensional vectors held by n distributed users while ensuring differential privacy (DP) of the users' vectors. The utility of DP-DME is measured by the mean squared error (MSE) between the estimate at the server and the true mean. DP-DME has been studied under multiple notions of DP. Local differential privacy (LDP) and distributed DP with secure aggregation (SecAgg) are the most common notions of DP used in DP-DME settings with an untrusted server. LDP provides strong resilience to dropouts, colluding users, and adversarial attacks, but suffers from poor utility. In contrast, SecAgg-based DP-DME achieves an $O(n)$ utility gain over LDP in DME, but requires increased communication and computation overheads and complex multi-round protocols to handle dropouts and attacks.

We ask: *How can we design a DP-DME mechanism that improves utility beyond LDP while avoiding the overheads of SecAgg?* Our key observation is that SecAgg achieves higher utility by using perfectly correlated noise, while LDP achieves resilience to attacks and dropouts by relying on independent noise. To unify these extremes, we propose a generalized framework for DP-DME in which LDP and SecAgg appear as boundary cases. In this framework, user i sends $\mathbf{Y}_i = \mathbf{x}_i + \mathbf{Z}_i$ to the central server where $\mathbf{x}_i$ is the private vector and $\mathbf{Z}_i$ is a random noise vector used for privacy. We then perform an information-theoretic analysis to find the optimum covariance structure among $\{\mathbf{Z}_i\}_{i=1}^n$ that minimizes MSE under any given (ϵ, δ)-DP budget, accounting for both user dropouts and collusion. Building on the results of this analysis, we introduce CorDP-DME, a novel single-round protocol based on the correlated Gaussian mechanism. By carefully optimizing the noise correlations across users, CorDP-DME lowers estimation error while preserving robustness to dropouts and collusion, achieving the best of both LDP and SecAgg-based methods. Our results show that CorDP-DME consistently outperforms LDP in utility and achieves significant improvements over SecAgg in terms of overheads, providing the first principled DP-DME mechanism that balances efficiency, robustness, and accuracy. Table I shows a comparison of CorDP-DME with existing methods across a number of features.

	LDP	**Distributed DP with SecAgg**	**CorDP-DME (ours)**
Rounds in protocol	Single round	Multiple rounds	Single round
MSE: no dropouts	$O\left(\frac{d}{n}\right)$	$O\left(\frac{d}{n^2}\right)$	$O\left(\frac{d}{n^2}\right)$
MSE: dropouts $\leq p$	$O\left(\frac{d}{n-p}\right)$	$O\left(\frac{d}{(n-p)^2}\right)$	$O\left(\frac{dp}{n(n-p)}\right)$
Computation	User: $O(d)$ Server: $O(d)$	User: $O(n^2+dn)$ Server: $O(dn^2)$	User: $O(dn)$ Server: $O(d)$
Dropouts	Gradual rise in MSE with dropouts	Gradual rise in MSE with dropouts up to the threshold, MSE surge afterwards	Gradual rise in MSE with dropouts
Collusion	(ϵ, δ)-DP with any number of colluding users	(ϵ, δ)-DP up to $c < n/3$ colluding users, sudden drop in privacy afterwards	(ϵ, δ)-DP up to *any* c colluding users, graceful privacy decay afterwards

TABLE I
COMPARISON OF CORDP-DME WITH EXISTING APPROACHES FOR DP-DME.