Numerical Accounting in the Shuffle Model of Differential Privacy

Antti Koskela *antti.h.koskela@nokia-bell-labs.com*
Nokia Bell Labs
University of Helsinki

Mikko Heikkilä *mikko.a.heikkila@helsinki.fi*
Department of Computer Science
University of Helsinki

Antti Honkela *antti.honkela@helsinki.fi*
Department of Computer Science
University of Helsinki

Reviewed on OpenReview: https://openreview.net/forum?id=11osftjEbF

Abstract

Shuffle model of differential privacy is a novel distributed privacy model based on a combination of local privacy mechanisms and a secure shuffler. It has been shown that the additional randomisation provided by the shuffler improves privacy bounds compared to the purely local mechanisms. Accounting tight bounds, however, is complicated by the complexity brought by the shuffler. The recently proposed numerical techniques for evaluating (ε, δ)-differential privacy guarantees have been shown to give tighter bounds than commonly used methods for compositions of various complex mechanisms. In this paper, we show how to utilise these numerical accountants for adaptive compositions of general ε-LDP shufflers and for shufflers of k-randomised response mechanisms, including their subsampled variants. This is enabled by an approximation that speeds up the evaluation of the corresponding privacy loss distribution from $\mathcal{O}(n^2)$ to $\mathcal{O}(n)$, where n is the number of users, without noticeable change in the resulting $\delta(\varepsilon)$-upper bounds. We also demonstrate looseness of the existing bounds and methods found in the literature, improving previous composition results for shufflers significantly.

1 Introduction

The shuffle model of differential privacy (DP) is a distributed privacy model which sits between the high trust–high utility centralised DP, and the low trust–low utility local DP (LDP). In the shuffle model, the individual results from local randomisers are only released through a secure shuffler. This additional randomisation leads to "amplification by shuffling", resulting in better privacy bounds against adversaries without access to the unshuffled local results.

We consider computing privacy bounds for both single and composite shuffle protocols, where by composite protocol we mean a protocol, where the subsequent user-wise local randomisers depend on the same local datasets and possibly on the previous output of the shuffler, and at each round the results from the local randomisers are independently shuffled. Moreover, using the analysis by Feldman et al. (2023), we provide bounds in the case the subsequent local randomisers are allowed to depend adaptively on the output of the previous ones.

In this paper we show how numerical accounting (Koskela et al., 2020; 2021; Gopi et al., 2021) can be employed for privacy analysis of both single and composite shuffle DP mechanisms. We demonstrate that thus obtained bounds can be up to orders of magnitudes tighter than the existing bounds from the literature. We also evaluate how significantly adversaries with varying capabilities differ in terms of the resulting privacy bounds using the k-randomised response mechanism. For conciseness, most of the proofs are given in the Appendix.

1.1 Related work

DP was originally defined in the central model assuming a trusted aggregator by Dwork et al. (2006), while the fully distributed LDP was formally introduced and analysed by Kasiviswanathan et al. (2011). Closely related to the shuffle model of DP, Bittau et al. (2017) proposed the Encode, Shuffle, Analyze framework for distributed learning, which uses the idea of secure shuffler for enhancing privacy. The shuffle model of DP was formally defined by Cheu et al. (2019), who also provided the first separation result showing that the shuffle model is strictly between the central and the local models of DP. Another direction initiated by Cheu et al. (2019) and continued, e.g., by Balle et al. (2020b); Ghazi et al. (2021) has established a separation between single- and multi-message shuffle protocols.

There exists several papers on privacy amplification by shuffling, some of which are central to this paper. Erlingsson et al. (2019) showed that the introduction of a secure shuffler amplifies the privacy guarantees against an adversary, who is not able to access the outputs from the local randomisers but only sees the shuffled output. Balle et al. (2019) improved the amplification results and introduced the idea of privacy blanket, which we also utilise in our analysis of k-randomised response. Feldman et al. (2021) used a related idea of hiding in the crowd to improve on the previous results, and their analysis was further improved in (Feldman et al., 2023). Girgis et al. (2021) generalised shuffling amplification further to scenarios with composite protocols and parties with more than one local sample under simultaneous communication and privacy restrictions. We use the improved results of Feldman et al. (2023) in the analysis of general LDP mechanisms, and compare our bounds with theirs in Section 3.3. We also calculate privacy bounds in the setting considered by Girgis et al. (2021), namely in the case where a subset of users sending contributions to the shufflers are sampled randomly. This can be seen as a subsampled mechanism and we are able to combine the analysis of Feldman et al. (2023), the privacy loss distribution related subsampling results of Zhu et al. (2022) and FFT accounting to obtain tighter (ε, δ)-bounds than Girgis et al. (2021), as shown in Section 3.4.

2 Background: numerical privacy accounting

Before analysing the shuffled mechanisms we introduce some required theory and notations. In particular, we use the privacy loss distribution formalism, which is based on finding the so-called dominating pairs of distributions for the given mechanisms. For more detailed presentations of the theory, we refer to Koskela et al. (2021); Gopi et al. (2021); Zhu et al. (2022).

2.1 Differential privacy and privacy loss distribution

An input dataset containing n data points is denoted as $X = (x_1, \ldots, x_n) \in \mathcal{X}^n$, where $x_i \in \mathcal{X}$, $1 \leq i \leq n$. We say $X, X' \in \mathcal{X}^n$ are neighbours if we get one by substituting one element in the other (denoted $X \sim X'$).

Definition 1. *Let $\varepsilon > 0$ and $\delta \in [0,1]$. Let P and Q be two random variables taking values in the same measurable space $\mathcal{O}$. We say that P and Q are (ε, δ)-indistinguishable, denoted $P \simeq_{(\varepsilon,\delta)} Q$, if for every measurable set $E \subset \mathcal{O}$ we have*

$$\Pr(P \in E) \leq \mathrm{e}^{\varepsilon} \Pr(Q \in E) + \delta, \qquad \Pr(Q \in E) \leq \mathrm{e}^{\varepsilon} \Pr(P \in E) + \delta.$$

Definition 2. *Let $\varepsilon > 0$ and $\delta \in [0,1]$. Mechanism $\mathcal{M} : \mathcal{X}^n \to \mathcal{O}$ is (ε, δ)-DP if for every $X \sim X'$: $\mathcal{M}(X) \simeq_{(\varepsilon,\delta)} \mathcal{M}(X')$. We call $\mathcal{M}$ tightly (ε, δ)-DP, if there does not exist $\delta' < \delta$ such that $\mathcal{M}$ is (ε, δ')-DP.*

When the data are distributed among several parties, and the local datasets are only accessed via purely local DP mechanisms, we say that the mechanisms guarantee local DP (LDP) and call the local DP mechanisms local randomisers (Kasiviswanathan et al., 2011).