Tight Differential Privacy Guarantees for the Shuffle Model with k-Randomized Response

Sayan Biswas[1,2,3], Kangsoo Jung[1(✉)], and Catuscia Palamidessi[1,2]

[1] Inria, Palaiseau, France
gangsoo.zeong@inria.fr
[2] École Polytechnique, Palaiseau, France
[3] EPFL, Lausanne, Switzerland

Abstract. Most differentially private algorithms assume a central model in which a reliable third party inserts noise to queries made on datasets, or a local model where the data owners directly perturb their data. However, the central model is vulnerable via a single point of failure, and the local model has the disadvantage that the utility of the data deteriorates significantly. The recently proposed shuffle model is an intermediate framework between the central and local paradigms. In the shuffle model, data owners send their locally privatized data to a server where messages are shuffled randomly, making it impossible to trace the link between a privatized message and the corresponding sender. In this paper, we theoretically derive the tightest known differential privacy guarantee for the shuffle models with k-Randomized Response (k-RR) local randomizers, under histogram queries, and we denoise the histogram produced by the shuffle model using the matrix inversion method to evaluate the utility of the privacy mechanism. We perform experiments on both synthetic and real data to compare the privacy-utility trade-off of the shuffle model with that of the central one privatized by adding the state-of-the-art Gaussian noise to each bin. We see that the difference in statistical utilities between the central and the shuffle models shows that they are almost comparable under the same level of differential privacy protection.

Keywords: Differential privacy · Shuffle model · Privacy-utility optimization

1 Introduction

As machine learning and data analysis using sensitive personal data are becoming more and more popular, concerns about privacy violations are also increasing manifold. The most successful approach to address this issue is differential privacy (DP). Most research performed in this area probes two main directions. One is the so-called central model, in which a trusted third party (the curator) collects

M. Mosbah et al. (Eds.): FPS 2023, LNCS 14551, pp. 440–458, 2024.
https://doi.org/10.1007/978-3-031-57537-2_27

the user's personal data and obfuscates them with a differentially private mechanism. The other is the local model, where the data owners apply the mechanism themselves on their data and send the perturbed data to the collector. A major drawback of the central model is that there is the risk that the curator may be corrupted. On the other hand, in the local model, there is no need to rely on a trusted curator. However, since each record is obfuscated individually, the utility of the data is substantially deteriorated compared to the central model.

In order to address the problem of the loss of utility in the local model, an intermediate paradigm between the central and the local models, known as the *shuffle model (SM)* of differential privacy, was recently proposed [5]. As an initial step, the shuffle model uses a local mechanism to perturb the data individually like the local model. The difference is that, after this first step of sanitization, a shuffler uniformly permutes the noisy data to dissolve their link with the corresponding data providers. Since a potential attacker is oblivious to the shuffling process, the data providers obtain two layers of privacy protection: injection of random noise by the local randomizer and anonymity by data shuffling. This allows the shuffle model to achieve a certain level of privacy protection using less noise than the local model.

The privacy guarantees provided by the shuffle model have been rigorously analyzed in several studies. More specifically, given a local mechanism with a level of privacy parameterised by ϵ_0 (pure local DP) or (ϵ_0, δ_0) (approximate local DP), the aim is to derive a (ϵ, δ) bound on the level of differential privacy guaranteed by applying shuffling on top of the local mechanism. In this paper, we derive the tight (ϵ, δ)-DP guarantee for the shuffle model with the k-RR local mechanism by using the concept of (ϵ, δ)-adaptive differential privacy (ADP) proposed by Sommer et al. in [15]. Next, we consider the question of how convenient the shuffle model is for publishing histograms in terms of the privacy-utility trade-off as opposed to the central model.

We perform various experiments on both synthetic and real data (the Gowalla dataset) and compare the utilities of the two models calibrated with the same privacy parameters. As expected, the utility of the central model is better than that of the shuffle model, consistent with what was observed in the literature [6]. However, in our case, the gap is very small – namely the histograms resulting from the shuffle model, once de-noised, are almost as close to the original ones as those of the central model. The contributions of this paper are as follows.

1. we derive an analytical form of the tight differential privacy guarantee for the shuffle model with k-RR local randomizer under histogram queries, and therefore, show that the shuffle model, essentially, provided a higher level of DP guarantee than what is known by the community, for the same level of locally injected noise to the data.
2. using the tight bound of the (ϵ, δ)-DP provided by the shuffle model, as derived, we compare the privacy-utility trade-off of the shuffle model and the optimized Gaussian mechanism for the histogram queries and show that their performances are comparable.

2 Related Work

Recently, intensive research on shuffle models of differential privacy has been done in various directions. One of the major research directions in this area is the study of privacy amplification by shuffling [3,10]. Erlingsson et al. [10] analysed the privacy amplification of the local randomizer's privacy protection by shuffling. Balle et al. [3] introduced the idea of privacy guarantee in shuffle models and quantitatively analyzed the relationship between the privacy parameter ϵ and the number of participants in the shuffle protocol. Feldman et al. [11] improved Balle et al.'s results and suggested an asymptotically optimal dependence of the privacy amplification on the privacy parameter of the local randomizer. However, neither [3] nor [11] explicitly theorize any guarantee for the tightness of the bounds for the privacy guarantee of shuffle models. Koskela et al. [14] proposed computational methods to estimate tight bounds based on weak adversaries – however, they are not expressed by an analytical formula, they can only be computed via an algorithm. Sommer et al. introduced the notion of adapted differential privacy (ADP) in [15] and laid down specific conditions to achieve the tight (ϵ, δ)-ADP for any abstract and high-level probabilistic mechanism. To derive the tight DP guarantees for SMs, we adapt Sommer et al.'s result and obtain necessary and sufficient conditions for achieving δ that warrants the best (ϵ, δ)- DP guarantee in SMs with a k-RR local randomizer.

3 Preliminaries

Definition 1 (Differential privacy [9]). *For a certain query, a randomizing mechanism $\mathcal{K}$ is* (ϵ, δ)-differentially private (DP) *if for all adjacent datasets, D_1 and D_2, and all $S \subseteq$ Range($\mathcal{K}$), we have:*

$$\mathbb{P}[\mathcal{K}(D_1) \in S] \leq e^{\epsilon}\, \mathbb{P}[\mathcal{K}(D_2) \in S] + \delta$$

Definition 2 (Adaptive differential privacy [15]). *For $x_0, x_1 \in \mathcal{X}$, where $\mathcal{X}$ is the space of the original data, and for a member u in the dataset, a randomizing mechanism $\mathcal{K}$ is* (ϵ, δ)-adaptive differentially private (ADP) for x_0 and x_1 *if for all datasets, $D(x_0)$ and $D(x_1)$, and all $S \subseteq$ Range($\mathcal{K}$), we have:*

$$\mathbb{P}[\mathcal{K}(D(x_0)) \in S] \leq e^{\epsilon}\, \mathbb{P}[\mathcal{K}(D(x_1) \in S] + \delta$$

where $D(x_0)$ and $D(x_1)$ are datasets differing only in the entry of the fixed member u: $D(x)$ means that u reports x for every $x \in \mathcal{X}$, keeping the entries of all the other users the same.

Remark 1. $\mathcal{K}$ is (ϵ, δ)-DP implies that $\mathcal{K}$ is (ϵ, δ)-ADP for every $x_0, x_1 \in \mathcal{X}$.

Definition 3 (Tight DP (or ADP) [15]). *Let $\mathcal{K}$ be (ϵ, δ)-DP (or ADP for $x_0, x_1 \in \mathcal{X}$). We say that δ is* tight *for $\mathcal{K}$ (w.r.t. ϵ and x_0, x_1 in case of ADP) if there is no $\delta' < \delta$ such that $\mathcal{K}$ is (ϵ, δ')-DP (or ADP for x_0, x_1).*

Definition 4 (Local differential privacy [8]). *Let $\mathcal{X}$ denote a possible alphabet for the original data and let $\mathcal{Y}$ be the alphabet of noisy data. A randomizing mechanism $\mathcal{R}$ provides* ϵ-local differential privacy (LDP) *if for all $x_1, x_2 \in \mathcal{X}$, and all $y \in \mathcal{Y}$, we have*

$$\mathbb{P}[\mathcal{R}(x_1) = y] \leq e^{\epsilon}\, \mathbb{P}\left(\mathcal{R}(x_2) = y\right)$$

Definition 5 (k-Randomized Response [13]). *Let $\mathcal{X}$ be a discrete alphabet of size k. Then* k-randomized response *(k-RR) mechanism, $\mathcal{R}_{kRR}$, is a locally differentially private mechanism that stochastically maps $\mathcal{X}$ onto itself (i.e., $\mathcal{Y} = \mathcal{X}$), given by*

$$\mathcal{R}_{kRR}(y|x) = \begin{cases} c\, e^{\epsilon} & , \text{ if } x = y \\ c, & , \text{ otherwise} \end{cases}$$

for any $x, y \in \mathcal{X}$, where $c = \frac{1}{e^{\epsilon}+k-1}$.

Definition 6 (Shuffle model [10]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be discrete alphabets for the original and the noisy data respectively. For any dataset of size $n \in \mathbb{N}$, the* shuffle model (SM) *is defined as $\mathcal{M} : \mathcal{X}^n \mapsto \mathcal{Y}^n$, $\mathcal{M} = \mathcal{S} \circ \mathcal{R}^n$, where*

- *$\mathcal{R} : \mathcal{X} \mapsto \mathcal{Y}$ is a local randomizer, stochastically mapping each element of the input dataset, sampled from $\mathcal{X}$, onto an element in $\mathcal{X}$, providing ϵ_0-local differential privacy.*
- *$\mathcal{S} : \mathcal{Y}^n \mapsto \mathcal{Y}^n$ is a shuffler that uniformly permutes the finite set of messages of size $n \in \mathbb{N}$, that it takes as an input.*

A SM can be perceived as having a sequence of messages going through the mechanism $\mathcal{M}$ and then coming out as the frequencies of each of the noisy messages, as the idea of the layer of "shuffling" is to randomize the noisy messages w.r.t. their corresponding senders by a random permutation. Let us call this particular brand of query on SM as the *histogram query*.

Definition 7 (Histogram query [2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be discrete alphabets for the original and the noisy data respectively. For any dataset of size $n \in \mathbb{N}$, the* histogram query *on SM, $\mathcal{M} : \mathcal{X}^n \mapsto \mathbb{R}^{+n}$, is defined as $\mathcal{M} = \mathcal{T} \circ \mathcal{R}^n$, where*

- *$\mathcal{R} : \mathcal{X} \mapsto \mathcal{Y}$ is a local randomizer providing ϵ_0-local differential privacy, as in Definition 6.*
- *$\mathcal{T} : \mathcal{Y}^n \mapsto \mathbb{R}^n$ is a function that gives the frequency of each message in finite set of messages of size $n \in \mathbb{N}$, that it takes as an input.*

In other words, if we have a dataset $D_{\mathcal{X}} = (x_1, \ldots, x_n) \in \mathcal{X}^n$, then $D_{\mathcal{Y}} = \mathcal{M}(D_{\mathcal{X}}) = \mathcal{T}((\mathcal{R}(x_1), \ldots, \mathcal{R}(x_n)) = (s_1, \ldots, s_n)$, where $s_i = n_i/n$ with n_i denoting the number of times $\mathcal{R}(x_i)$ occurs in $D_{\mathcal{Y}}$.

Definition 8 (Privacy loss random variable [15]). *For a probabilistic mechanism mapping messages from the alphabet of original messages to the alphabet for noisy messages, $M : \mathcal{X} \mapsto \mathcal{Y}$, let us fix $x_0, x_1 \in \mathcal{X}$ and a potential output*

$y \in \mathcal{Y}$. *The* privacy loss random variable *of* y *for* x_0 *over* x_1 *is defined as: where* $M(x_i)$ *is the probability distribution of the noisy output for the original input* x_i *for* $i \in \{0, 1\}$.

$$\mathcal{L}_{M(x_0)/M(x_1)}(y) = \begin{cases} +\infty & \begin{cases} \mathbb{P}(M(x_0) = y) \neq 0, \\ \mathbb{P}(M(x_1) = y) = 0 \end{cases} \\ \ln \frac{\mathbb{P}(M(x_0)=y)}{\mathbb{P}(M(x_1)=y)} & \begin{cases} \mathbb{P}(M(x_0) = y) \neq 0, \\ \mathbb{P}(M(x_1) = y) \neq 0 \end{cases} \\ -\infty & o.w. \end{cases} \tag{1}$$

Definition 9 (Privacy loss distribution [15]). *Let* P_1 *and* P_2 *be two probability distributions on* $\mathcal{Y}$ *(the finite alphabet for noisy messages). The* privacy loss distribution, ω, for A over B *is defined as:*

$$\omega(u) = \sum_{y:\mathcal{L}_{A/B}(y)=u} \mathbb{P}(A = y) \text{ for all } u \in \mathcal{U} \text{ , where } \mathcal{U} = \bigcup_{y \in \mathcal{Y}} \{\mathcal{L}_{A/B}(y)\} \subset \mathbb{R} \text{ .}$$

4 Tight Privacy Guarantee for SM

4.1 Overview

Sommer et al. in [15] proposed a notion of adaptive differential privacy and derived a very important sufficient and necessary result for any probabilistic mechanism to have the best formal privacy guarantee. Adaptive differential privacy essentially translates the idea of a differential privacy guarantee with respect to a chosen pair of elements in the dataset. Exploiting this result (Result 1), we derived the necessary and sufficient condition needed to warrant the best DP guarantee for SM with the most popularized LDP satisfying local randomizer, the k-RR mechanism. This essentially draws the tight DP guarantee that an SM can induce being locally randomized with a k-RR mechanism. At the crux of this paper, the importance of deriving the tight DP guarantee by SM under the k-RR local randomizer implies that we show that the SM provides a higher level of privacy than what is known by the existing work in the literature that focuses on improving the privacy bound for the SM.

Table 1. Value of δ derived from the existing work and our proposed result

	[10]	[3]	[11]	[14]	Proposed tight δ
$\epsilon = 0.1$	0.97	0.229	0.066	9.01E-4	2.38E-28
$\epsilon = 0.2$	0.89	0.002	1.91E-5	1.89E-6	1.61E-42
$\epsilon = 0.3$	0.77	1.77E-6	2.43E-11	2.19E-10	5.22E-57
$\epsilon = 0.4$	0.64	5.95E-11	1.35E-19	3.14E-16	5.14E-72

Table 1 presents the values of δ obtained from the results in [3,10,11,14] and the proposed derivation in (6) of this paper, by varying ϵ from 0.1 to 0.4, fixing $n = 100$ and $\epsilon_0 = 0.5$. We observe that, indeed, the value of δ computed from (6) in Definition 11 is significantly less compared to the other existing improvements proposed, highlighting that our proposed result engenders the best possible DP guarantee for SMs under the k-RR local randomizer.

4.2 Framework

Let $\mathcal{X} = (x_0, \ldots, x_{k-1})$ be the alphabet of messages of size $k \in \mathbb{N}$, $k > 1$ and $\mathfrak{U}$ be the set of all users involved in the environment. For simplicity, we assume the alphabets of the original and noisy messages to be the same, both being $\mathcal{X}$. Therefore, the local randomizer of our shuffle mechanisms locally sanitizes the dataset by mapping original messages sampled from $\mathcal{X}$ to elements of $\mathcal{X}$.

Let ϵ_0 be the privacy parameter of $\mathcal{R}_{\text{kRR}}$, which is used as the local randomizer for the shuffle mechanisms discussed in this paper. Furthermore, letting $D_{\mathcal{X}}$ be the dataset of the original messages of n users, each of which is sampled from (and obfuscated to) $\mathcal{X}$, we denote $D_{\mathcal{X}\,z}$ as the original message of $z \in \mathfrak{U}$ in $D_{\mathcal{X}}$ for any $z \in \mathfrak{U}$. Let $D_{\mathcal{Y}} = \mathcal{R}^n_{\text{kRR}}(D_{\mathcal{X}}) = \{\mathcal{R}_{\text{kRR}}(D_{\mathcal{X}\,z}) : z \in \mathfrak{U}\}$ be the noisy dataset going through $\mathcal{R}_{\text{kRR}}$.

For the purpose of analysing the adaptive differential privacy, let us fix a certain user, $u \in \mathfrak{U}$, whose data is in $D_{\mathcal{X}}$. Since the only major distinction that k-RR mechanism makes in the process of mapping a datum from its original value to the obfuscated value is whether the original value and the obfuscated value are the same or not (i.e., the probability that the x is being reported as x' is the same for every $x' \in \mathcal{X}$ when $x \neq x'$), it is reasonable for us to study the adaptive differential privacy guarantee with respect to a couple of potential original messages of u, say $x_0, x_1 \in \mathcal{X}$, $x_0 \neq x_1$ in the environment where the shuffle model uses a k-RR local randomizer.

The idea behind adaptive differential privacy w.r.t. x_0, x_1 is to make it significantly difficult to predict whether u's original message is x_0 or x_1. In the context of this work, since we will be focusing on the case of having the local randomizer as the k-RR mechanism, the only gravity x_1 holds as far as the shuffle model is concerned is the fact that it is different from x_0. Thus x_1 could represent any $x \in \mathcal{X}$ such that $x \neq x_0$. Therefore, we shall be analysing the privacy of u's original message being x_0 and compare its privacy level of being identifiable with a different potential original message, which we fix as x_1 w.l.o.g. Let's call x_0 as the *primary input* for u and x_1 be the *secondary input*. For a fixed set of values reported by every user in $\mathfrak{U} \setminus \{u\}$, let $D(x_0)$ represent the edition of the dataset where u reports x_0, and let $D(x_1)$ represent the one where u reports x_1.

The most important result from literature - Lemma 5 in [15] - that is heavily exploited in this paper is as follows:

Result 1: (Lemma 5 [15]) For every probabilistic mechanism $M : \mathcal{X} \mapsto \mathcal{Y}$, for any $x_0, x_1 \in \mathcal{X}$ and any ϵ, $\delta(\epsilon) > 0$, M is tightly (ϵ, δ)-ADP for x_0, x_1 iff

$$\delta(\epsilon) = \omega(\infty) + \sum_{\substack{u \in \mathcal{U} \setminus \{\infty, -\infty\} \\ u > \epsilon}} (1 - e^{\epsilon - u})\omega(u) \tag{2}$$

4.3 Theorems and Results

As we are interested in finding $\epsilon > 0$ and, correspondingly, $\delta > 0$ that provide a tight ADP guarantee for $\mathcal{M}$ for x_0, x_1, we define the constants $\kappa_1, \kappa_2, \kappa_3$ to simplify the mathematical results derived in the subsequent sections as follows:

$$\kappa_1 := \frac{e^{\epsilon_0}(e^{\epsilon_0} + k - 2)}{k - 1} \tag{3}$$

$$\kappa_2 := \frac{k - 1}{e^{\epsilon_0} + k - 2} \tag{4}$$

$$\kappa_3 := \frac{(k-1)^{n_{x_0}}(e^{\epsilon_0} + k - 2)^{n - n_{x_0} - s}}{(e^{\epsilon_0} + k - 1)^n} \tag{5}$$

Remark 2. Note that $\kappa_1, \kappa_2, \kappa_3 > 0$ for any $\epsilon_0 > 0$, $n \in \mathbb{N}$, $k \in \mathbb{N}_{\geq 2}$, $s \in \mathbb{N}$.

From now on we shall focus on the histogram query of the shuffle model. For the same ϵ_0-LDP mechanism $\mathcal{R}_{\text{kRR}}$ to be used as the local randomizer for histogram query, let $\mathcal{M} = \mathcal{T} \circ \mathcal{R}_{\text{kRR}}$ denote the shuffle model that takes in a sequence of original messages, obfuscates them locally using $\mathcal{R}_{\text{kRR}}$, and broadcasts the frequency of each message in the noisy dataset. In other words, having u having x_i as her original message for $i \in \{0, 1\}$, $\mathcal{M}(D(x_i)) = (\mathcal{M}_{x_0}(x_i), \ldots, \mathcal{M}_{x_{k-1}}(x_i))$ where $\mathcal{M}_{x_j}(x_i)$ is a random variable giving the frequency of $x_j \in \mathcal{X}$ in the noisy dataset, $D_{\mathcal{Y}}$, obfuscated by $\mathcal{R}_{\text{kRR}}$. Assuming that u's original data is x_0 (w.l.o.g.), let n_{x_0} denote the number of times x_0 has appeared in $D_{\mathcal{X}}$ for the original entries from all users in $\mathfrak{U} \setminus u$.

Definition 10. *By Definition 8, the* privacy loss random variable for the histogram query for shuffle model *of x_0 over x_1 with respect to a certain output $s \in \mathbb{N}$, in $\mathcal{M}$ is $v_s(x_0, x_1) = \ln \frac{\mathbb{P}(\mathcal{M}_{x_0}(x_0) = s)}{\mathbb{P}(\mathcal{M}_{x_0}(x_1) = s)}$.*

Definition 11. *For $s \in \{0, \ldots, n\}$, $r \in \{0, \ldots, s\}$, let $\mu(s, r) = \binom{n_{x_0}}{r}\binom{n - n_{x_0}}{s - r}\kappa_1^r$ and $\tau_r = \kappa_2(n - n_{x_0}) + (e^{\epsilon_0} - \kappa_2)(s - r)$. For any $\epsilon > 0$, let us define*

$$\hat{\delta}(\epsilon) := \sum_{s=0}^{n} \mathbb{1}_{\{v_s(x_0, x_1) > \epsilon\}}(1 - e^{\epsilon - v_s(x_0, x_1)})\frac{\kappa_3}{n - n_{x_0}} \sum_{r=0}^{s} \mu(s, r)\tau_r \tag{6}$$

where $\mathbb{1}_E$ is the indicator function for any event E.

Theorem 1. *For any $\epsilon > 0$, we get the tight (ϵ, δ)-ADP guarantee for $\mathcal{M}$ with respect to x_0, x_1 iff $\delta = \hat{\delta}(\epsilon)$ as in as in (6) of Definition 11 where*

$$v_s(x_0, x_1) = \ln\left(\kappa_2 + \frac{\frac{(e^{\epsilon_0}-\kappa_2)}{n-n_{x_0}}\left(\sum_{r=0}^{s}(s-r)\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r\right)}{\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-n_{x_0}}{s-r}\kappa_1^r}\right).$$

Corollary 1. *For any $\epsilon > 0$, we get the tight (ϵ, δ)-DP guarantee for $\mathcal{M}$ iff:*

$$\delta(\epsilon) \coloneqq \sum_{s=0}^{n} \mathbb{1}_{\{v_s > \epsilon\}}(1 - e^{\epsilon - v_s})\frac{\kappa_3}{n - n_{x_0}}\sum_{r=0}^{s}\mu(s, r)\tau_r \tag{7}$$

where $v_s = \max_{x_0, x_1 \in \mathcal{X}} v_s(x_0, x_1)$ and $v_s(x_0, x_1)$ is as derived in Theorem 1.

5 Evaluating the Utility of the Shuffle Model

It is crucial to have the tight bound in the privacy guarantee for shuffle models to be able to conduct a fair comparison of utilities of shuffle models with other forms of differential privacy under a certain level of privacy protection.

Suppose with ϵ, δ, we get a tight (ϵ, δ)-ADP guarantee for $\mathcal{M}$ w.r.t. x_0 as the primary input. We wish to compare how the utility of $\mathcal{M}$ would perform against that of a central model of differential privacy for histogram query implemented on $D_{\mathcal{X}}$ with the same privacy parameters ϵ and δ. For this, we will be sticking to the most optimal framework, known until now [4], of one of the most popular mechanisms for the central model for (ϵ, δ)-DP: the *Gaussian mechanism*. The details of the theoretical build-up are provided in Appendix B.

In [7], Cheu et al. give theoretical evidence that the accuracy of the SM lies in between the central and local models of DP. However, no experimental analysis had been performed to dissect how low the accuracy of SMs lies when compared to the central model when both provide the same level of privacy protection. Thus, the main goal of our experiments was to empirically show the scale of difference in accuracy between SM and the central model by comparing their statistical utilities under the tight and equal DP guarantee. To do this end, we compared the statistical approximation of the true distribution from the SM with k-RR local randomizer to that of the central model by applying the Gaussian mechanism [4], using the value of δ derived from (6), ensuring the tight (ϵ, δ)-DP guarantee.

5.1 Experimental Results on Synthetic Data

In this section, we carry out an experimental analysis to illustrate the comparison of utilities for histogram query of the shuffle model using k-RR local randomizer and the optimal Gaussian mechanism using synthetically generated data sampled from $\mathcal{N}(0, 2)$. We experimented and demonstrated our results in

the two categories: (i) trend analysis of δ providing the tight ADP guarantee for $\mathcal{M}$ and (ii) utility comparison between $\mathfrak{N}$ and $\mathcal{M}$ under the same level of differential privacy.

To analyze the values of δ providing a tight ADP guarantee for $\mathcal{M}$, we change the values of ϵ, ϵ_0, n, n_0, and k that enable us to see the change in the trend of δ. For comparing the utilities of the central model and the shuffle model, we considered $\hat{\delta}$ as in (15), providing the worst possible tight ADP over every $x_0, x_1 \in \mathcal{X}$, and therefore, by Remark 1, a DP guarantee. Table 2 shows the default values of the parameters used for the experiment.

Table 2. Experimental parameters used for synthetic data

Parameter name	Values
ϵ	0.1 to 3
ϵ_0	0.1 to 3
n	50, 100, 150, 1000, 100000
x_0	1 to 15
k	5, 10, 15

Tight δ for Histogram Queries. We show the experimental results for deriving δ providing (ϵ, δ)-ADP guarantee, as given by Theorem 1, by changing the values for ϵ, ϵ_0, n and k. We use the *total variation distance*, $d_{TV}(.)$, to evaluate $\mathcal{W}(\mathcal{M})$ and $\mathcal{W}(\mathfrak{N})$ – the "distances" of the estimated original distribution obtained from shuffle model with k-RR local randomizer, using matrix inversion, (shuffle+INV), and the distribution sanitized with Gaussian mechanism from the original distribution itself. Table 3 shows δ when we vary ϵ, for three categories:

(a) We change ϵ_0, fixing $n_{x_0} = 80$, $n = 100$, and $k = 10$. We observe that δ decreases as ϵ increases for the same ϵ_0, and δ increases as ϵ_0 increases under a fixed value of ϵ. When it does not satisfy the $v_s > \epsilon$ condition of equation (57), δ becomes 0. For a fixed ϵ and ϵ_0, a high value of δ decreases the level of privacy protection. Thus, experimentally, we can validate that for a constant ϵ, δ increases as ϵ_0 used for k-RR increases, ensuring that the privacy protection of the shuffle model decreases with a decrease in the privacy level of its local randomizer.

(b) We vary n fixing $k = 10$, $\epsilon_0 = 2$, and $n_{x_0} = 80$. For the same ϵ, δ becomes smaller as the value of n increases. A lower δ means higher privacy protection, reassuring that the shuffle model provides higher privacy protection as the number of users (samples) increases.

(c) We alter k fixing $n = 100$, $\epsilon_0 = 2$, and $n_{x_0} = 80$. As the value of k increases, δ decreases. This is also due to the characteristic of the k-RR mechanism, which is used as the local randomizer for $\mathcal{M}$. The inference probability for a potential adversary decreases as the size of the domain for the data increases.

Table 3. Tight δ for different ϵ

Varying ϵ_0							
	$\epsilon = 0.1$	$\epsilon = 0.5$	$\epsilon = 1.0$	$\epsilon = 1.5$	$\epsilon = 2.0$	$\epsilon = 2.5$	$\epsilon = 3.0$
$\epsilon_0 = 1$	2.08E-20	3.42E-43	0	0	0	0	0
$\epsilon_0 = 2$	2.49E-15	3.25E-22	2.20E-30	1.57E-40	0	0	0
$\epsilon_0 = 3$	8.79E-11	5.73E-13	3.52E-16	4.49E-20	1.09E-25	4.52E-33	0
Varying n							
n =50	1.91E-08	5.02E-12	2.40E-15	4.49E-21	0	0	0
n =100	2.49E-15	3.25E-22	2.20E-30	1.57E-40	0	0	0
n =150	6.58E-22	7.83E-32	2.75E-44	6.99E-59	0	0	0
Varying k							
k =5	1.96E-10	2.51E-14	1.08E-19	2.02E-27	0	0	0
k =10	2.49E-15	3.25E-22	2.20E-30	1.57E-40	0	0	0
k =15	1.66E-18	7.13E-28	1.49E-38	7.35E-50	0	0	0

Comparing the Utility of the Shuffle and the Central Models. In this section, we compare the utilities of the central model and the shuffle models, providing the same level of privacy protection. For neutral comparison, we perform the experiments into two cases: individual specific utility and community level utility, as described in Appendix B. We use the *total variation distance* to estimate the difference between original distribution and estimated distributions.

Table 4. Individual specific utility comparison of central and shuffle models for synthetic data ($\epsilon = 4$)

x_0	$n = 1,000$		$n = 100,000$	
	Gaussian	shuffle+INV	Gaussian	shuffle+INV
1	3E-3	1E-3	6E-6	3E-3
3	6E-4	2E-4	1E-5	5E-4
5	12E-4	11E-4	1E-5	5E-4
7	1E-4	4E-3	8E-6	3E-5

Table 4 shows the results from the experimental analysis of comparing the individual specific utilities of $\mathcal{M}$ and $\mathfrak{N}$ as the primary input, x_0, is changed. We performed the experiments for the case of $n = 1,000$ and $n = 100,000$, setting $\epsilon_0 = 4$, $\epsilon = 4$, and $k = 15$, calculating δ for each x_0. When $n = 1,000$, shuffle+INV is comparable with the Gaussian mechanism, depending on the value of x_0. However, when n is 100,000, the Gaussian mechanism shows better results regardless of x_0. This is explained through our choice of δ (given by Theorem 1), which depends on n_{x_0}, which, in turn, varies with x_0 and that

Gaussian mechanism inserts fixed noise regardless of n. However, even for a large value of n, the utility of shuffle+INV, although slightly worse than the Gaussian mechanism, is quite good as $\overline{\mathcal{W}}(\mathcal{M}, x_0)$ remains very low across different x_0.

For the community level utility, we apply the worst case (highest value) of δ computed over all the primary inputs for all the users in $\mathfrak{U}$, given as $\hat{\delta}$ in (15), to sanitize all input messages of the dataset – thus establishing the worst tight ADP guarantee possible on the shuffle model. This is used to determine the community level utility of the corresponding shuffle model with the estimated differential privacy guarantee. Similar to the case of individual specific utility, experiments were performed for the case of $n = 1,000$ and $n = 100,000$, and the other parameters used for the experiment being the same. The experiments results are similar to what we showed for individual specific utility. When n is small, the utility of shuffle model is almost as much as that of the central model. As n increases, the utility of the Gaussian mechanism, $\mathfrak{N}$, improves slightly over that of the shuffle model under the same level of differential privacy, however they still are fairly close (Fig. 1).

5.2 Experimental Results on Real Data

Now we focus on the experimental results obtained using real location data from the Gowalla dataset [12]. Figure 2 illustrate the estimations of the original distributions of location data from San Francisco and Paris, respectively. We sanitize the original distribution using the shuffle model giving a tight differential privacy guarantee with parameters ϵ and $\hat{\delta}$, as in (15). We use the same ϵ and $\hat{\delta}$ to privatize the original data using the Gaussian mechanism as same in the previous experiment, thus getting a $(\epsilon, \hat{\delta})$-DP guarantee for both cases.

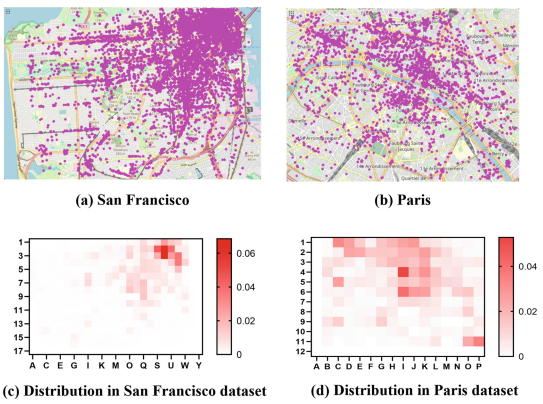

Fig. 1. (a) and (b): Location data from Gowalla check-ins from a northern part of San Francisco and a part of Paris. (c) and (d) give the heatmap of the locations in the areas of San Francisco and Paris as an alternative visualization.

To compare the utility of the two mechanisms under the same privacy level, we estimate the original distributions using shuffle+INV for the shuffle model and the Gaussian mechanism itself for the central model, as described in (13) and (14) and evaluate how far the corresponding estimations lie from the original distributions. We observe that the Gaussian mechanism approximates the original distributions slightly better than the shuffle+INV, but they are comparable.

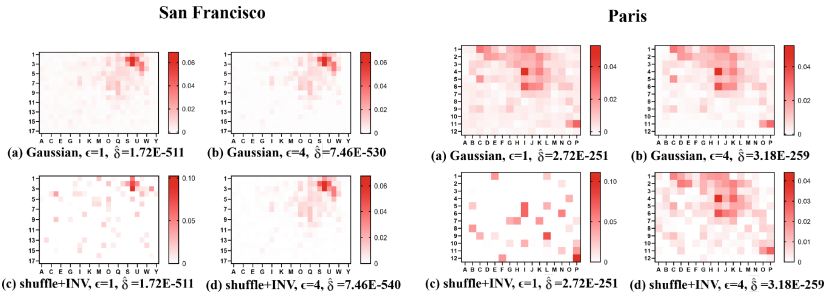

Fig. 2. Estimation of the original distribution from the noisy data obfuscated with the Gaussian mechanism and the SM in San Francisco and Paris dataset

As we observe in the previous experiment results, the number of samples, ϵ affects the utility. In Fig. 3, we show how the number of samples and the differential privacy parameters affect the utilities in more detail. In summary, we observe a consistency with the existing work in the trend of the Gaussian mechanism having a better utility than the shuffle model across all settings. However, when the number of samples is small and the privacy level is low, the utilities of the shuffle model and the central model are comparable.

Figure 3 (a) and (b) illustrate the evaluation of the TV distance between the original and the estimated distributions for San Francisco dataset. n ranges from $10,000$ to $100,000$, which is used to sample locations from the aforementioned San Francisco region. We set $\epsilon = 4$ and $\epsilon = 6$ to capture the change of distance between the original and the estimated distributions by varying n. We use $\hat{\delta}$, as in (15), to calculate community-level utility and we run the mechanism 10 times to obtain the boxplots. The results exhibit that shuffle model, $\mathcal{M}$, gives worse utility than the central model $\mathfrak{N}_{\epsilon,\hat{\delta}}$, and shuffle+INV shows better utility than shuffle. This trend is harmonious across the different settings for ϵ. It is reassuring to observe that the shuffle+INV is slightly closer or comparable with the Gaussian mechanism especially when the value of n is small ($n = 10,000$) and the privacy level is low ($\epsilon = 6$). Figure 3 (c) and (d) shows the TV distance between the estimated and the original distributions and the utility difference for locations in Paris dataset with n ranging from 1,000 to 10,000 and the other parameters being the same as the experiments for the San Francisco dataset.

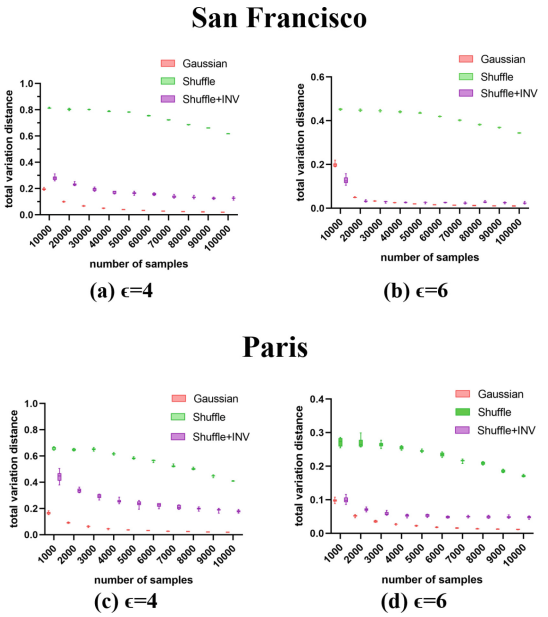

Fig. 3. Illustrating the comparison of community level utilities between Gaussian, shuffle and shuffle+INV for varying n and ϵ in San Francisco and Paris dataset

The overall trend of TV distance for the dataset of Paris is the same as that of San Francisco. Again, we observe that the utility of the shuffle+INV is better than that of just shuffle with k-RR, and the utilities of the shuffle+INV and the optimal Gaussian mechanism are almost indistinguishable when the number of samples and the privacy level are low. As we see from the heatmaps in Fig. 2, when the value of ϵ is 4, both the Gaussian mechanism and shuffle+INV generate results very close to the original distribution. Individual-specific utilities for the Paris and San Francisco datasets are described in Table 5.

Table 5. Individual specific utility comparison of central and shuffle models for Gowalla data ($\epsilon = 4$, $\epsilon_0 = 4$)

x_0	San Francisco		x_0	Paris	
	Gaussian	shuffle+INV		Gaussian	shuffle+INV
40	4E-6	1E-3	20	2E-6	3E-4
80	3E-5	5E-4	40	3E-5	2E-3
120	9E-6	1E-3	60	4E-5	2E-3
160	4E-5	2E-4	80	5E-5	4E-4
200	2E-5	2E-4	100	7E-5	1E-4

6 Conclusion

In this paper, we have compared the privacy-utility trade-off of two different models of differential privacy for histogram queries: the classic central model with the optimal Gaussian mechanism and the shuffle model with k-RR mechanism as the local randomizer, enhanced with post-processing to de-noise the resulting histogram. In order to do this comparison, we needed to derive the tight bounds for the level of privacy provided by the shuffle model, so that we could tune the parameters of the Gaussian mechanism to provide the same privacy.

First, we have used a result on the condition for tightness of ADP given by Sommer et al. in [15] and translated it in the context of shuffle models, giving rise to a closed form expression of the least δ for any ϵ and, thus, we obtained a necessary and sufficient condition to have the tight DP guarantee for the shuffle models. This result shows that the differential privacy ensured by the shuffle models under a certain level of local noise is much higher than what has been known by the community so far. Then, we performed experiments on synthetic and real location data from San Francisco and Paris, and we compared the statistical utilities of the shuffle and the central models. We observed that, although the central model still performs better than the shuffle model, only ever so slightly – the gap between their statistical utilities is very small and tends to vanish as the number of samples is small.

Acknowledgment. The work is supported by the European Research Council (ERC) project HYPATIA under the European Unions Horizon 2020 research and innovation programme. Grant agreement no. 835294 and ELSA - European Lighthouse on Secure and Safe AI funded by the European Union under grant agreement No. 101070617.

A Proof of Theorem Theorem 1

Setting $p = \mathbb{P}[x_0|x_0]$, $\overline{p} = \mathbb{P}[x_0|y \neq x_0]$ in $\mathcal{R}_{\text{kRR}}$, $\forall s \in [n]$, $\mathbb{P}[\mathcal{M}_{x_0}(x_0) = s]$

$$
\begin{aligned}
&= p \sum_{r=0}^{s-1} \left[\binom{n_{x_0}}{r} p^r (1-p)^{n_{x_0}-r} \binom{n-1-n_{x_0}}{s-1-r} \overline{p}^{s-1-r} (1-\overline{p})^{n-n_{x_0}-s+r} \right] \\
&+ (1-p) \sum_{r=0}^{s} \left[\binom{n_{x_0}}{r} p^r (1-p)^{n_{x_0}-r} \binom{n-1-n_{x_0}}{s-r} \overline{p}^{s-r} (1-\overline{p})^{n-n_{x_0}-1-s+r} \right] \\
&= \frac{e^{\epsilon_0}}{e^{\epsilon_0}+k-1} \sum_{r=0}^{s-1} \left[\binom{n_{x_0}}{r} \frac{e^{r\epsilon_0}(k-1)^{n_{x_0}-r}}{(e^{\epsilon_0}+k-1)^{n_{x_0}}} \binom{n-1-n_{x_0}}{s-1-r} \frac{(e^{\epsilon_0}+k-2)^{n-n_{x_0}-s+r}}{(e^{\epsilon_0}+k-1)^{n-1-n_{x_0}}} \right] \\
&+ \frac{k-1}{e^{\epsilon_0}+k-1} \sum_{r=0}^{s} \left[\binom{n_{x_0}}{r} \frac{e^{r\epsilon_0}(k-1)^{n_{x_0}-r}}{(e^{\epsilon_0}+k-1)^{n_{x_0}}} \binom{n-1-n_{x_0}}{s-r} \frac{(e^{\epsilon_0}+k-2)^{n-n_{x_0}-1-s+r}}{(e^{\epsilon_0}+k-1)^{n-1-n_{x_0}}} \right] \\
&= \frac{e^{\epsilon_0}(k-1)^{n_{x_0}}(e^{\epsilon_0}+k-2)^{n-n_{x_0}-s}}{(e^{\epsilon_0}+k-1)^n} \sum_{r=0}^{s-1} \binom{n_{x_0}}{r} \binom{n-1-n_{x_0}}{s-1-r} \kappa_1^r
\end{aligned}
$$

$$+\frac{(k-1)^{n_{x_0}+1}(e^{\epsilon_0}+k-2)^{n-n_{x_0}-1-s}}{(e^{\epsilon_0}+k-1)^n}\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-r}\kappa_1^r$$

$$=\kappa_3\left[e^{\epsilon_0}\sum_{r=0}^{s-1}\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r+\kappa_2\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-r}\kappa_1^r\right]$$

Using elementary combinatorial identities, we reduce to:

$$\kappa_3\left[\kappa_2\sum_{r=0}^{s}\binom{n_{x_0}}{r}\kappa_1^r\left(\binom{n-1-n_{x_0}}{s-1-r}+\binom{n-1-n_{x_0}}{s-r}\right)\right.$$

$$\left.+(e^{\epsilon_0}-\kappa_2)\left(e^{\epsilon_0}\sum_{r=0}^{s-1}\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r\right)\right]$$

$$=\kappa_3\left[\kappa_2\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-n_{x_0}}{s-r}\kappa_1^r+(e^{\epsilon_0}-\kappa_2)\left(\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r\right)\right]$$

$$=\kappa_3\left[\kappa_2\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-n_{x_0}}{s-r}\kappa_1^r+\frac{(e^{\epsilon_0}-\kappa_2)(s-r)}{n-n_{x_0}}\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-n_{x_0}}{s-r}\kappa_1^r\right]$$

$$=\frac{\kappa_3}{n-n_{x_0}}\sum_{r=0}^{s}\mu(s,r)\tau_r\ [\mu \text{ and } \tau \text{ are as in Definition 11}] \tag{8}$$

By similar arguments as above, for any $s\in\{0,\ldots,n\}$, $\mathbb{P}[\mathcal{M}_{x_0}(x_1)=s]$

$$=\frac{1}{e^{\epsilon_0}+k-1}\sum_{r=0}^{s-1}\left[\binom{n_{x_0}}{r}\frac{e^{r\epsilon_0}(k-1)^{n_{x_0}-r}}{(e^{\epsilon_0}+k-1)^{n_{x_0}}}\binom{n-1-n_{x_0}}{s-r}\frac{(e^{\epsilon_0}+k-2)^{n-n_{x_0}-s+r}}{(e^{\epsilon_0}+k-1)^{n-1-n_{x_0}}}\right]$$

$$+\frac{e^{\epsilon_0}+k-2}{e^{\epsilon_0}+k-1}\sum_{r=0}^{s}\left[\binom{n_{x_0}}{r}\frac{e^{r\epsilon_0}(k-1)^{n_{x_0}-r}}{(e^{\epsilon_0}+k-1)^{n_{x_0}}}\binom{n-1-n_{x_0}}{s-1-r}\frac{(e^{\epsilon_0}+k-1)^{n-n_{x_0}-1-s+r}}{(e^{\epsilon_0}+k-1)^{n-1-n_{x_0}}}\right]$$

$$=\kappa_3\left(\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r+\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-r}\kappa_1^r\right)$$

$$=\kappa_3\sum_{r=0}^{s}\binom{n_{x_0}}{r}\binom{n-n_{x_0}}{s-r}\kappa_1^r \tag{9}$$

Using Result 1, for every $k>2$ and $s\in\{0,1,\ldots,n\}$, we can say that $\mathcal{M}$ induces a tight (ϵ,δ)-ADP guarantee with respect to x_0, $x_1\in\mathcal{X}$ for any $\epsilon>0$ and δ iff δ is defined as:

$$\delta(\epsilon)=\sum_{v:v>\epsilon}(1-e^{\epsilon-v})\sum_{\substack{s=0\\ v=\ln\frac{\mathbb{P}[\mathcal{M}_{x_0}(x_0)=s]}{\mathbb{P}[\mathcal{M}_{x_0}(x_1)=s]}}}^{n}\mathbb{P}[\mathcal{M}_{x_0}(x_0)=s] \tag{10}$$

Using the expressions derived for $\mathbb{P}[\mathcal{M}_{x_0}(x_0)=s]$ and $\mathbb{P}[\mathcal{M}_{x_0}(x_1)=s]$ in (8) and (9), respectively, to get v_s:

$$= \ln \frac{\mathbb{P}[\mathcal{M}_{x_0}(x_0) = s]}{\mathbb{P}[\mathcal{M}_{x_0}(x_1) = s]} = \ln \frac{e^{\epsilon_0} \sum_{r=0}^{s-1} \binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r + \kappa_2 \sum_{r=0}^{s} \binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-r}\kappa_1^r}{\sum_{r=0}^{s-1} \binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r + \sum_{r=0}^{s} \binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-r}\kappa_1^r}$$

$$= \ln \left(\kappa_2 + \frac{(e^{\epsilon_0} - \kappa_2)\left(\sum_{r=0}^{s-1} \binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r\right)}{\sum_{r=0}^{s} \binom{n_{x_0}}{r}\binom{n-n_{x_0}}{s-r}\kappa_1^r} \right)$$

$$= \ln \left(\kappa_2 + \frac{\frac{(e^{\epsilon_0} - \kappa_2)}{n-n_{x_0}}\left(\sum_{r=0}^{s} (s-r)\binom{n_{x_0}}{r}\binom{n-1-n_{x_0}}{s-1-r}\kappa_1^r\right)}{\sum_{r=0}^{s} \binom{n_{x_0}}{r}\binom{n-n_{x_0}}{s-r}\kappa_1^r} \right) \tag{11}$$

Combining (10) and (11), $\delta(\epsilon) = \sum_{\substack{u:u>\epsilon; s=0 \\ v=\ln \frac{\mathbb{P}[\mathcal{M}_{x_0}(x_0)=s]}{\mathbb{P}[\mathcal{M}_{x_0}(x_1)=s]}}}^{n} (1 - e^{\epsilon - v})\mathbb{P}[\mathcal{M}_{x_0}(x_0) = s]$

$$= \sum_{s=0}^{n} \mathbb{1}_{\{v_s > \epsilon\}}(1 - e^{\epsilon - v_s})\mathbb{P}[\mathcal{M}_{x_0}(x_0) = s]$$

$$= \sum_{s=0}^{n} \mathbb{1}_{\{v_s > \epsilon\}}(1 - e^{\epsilon - v_s})\frac{\kappa_3}{n - n_{x_0}} \sum_{r=0}^{s} \mu(s, r)\tau_r = \hat{\delta}(\epsilon)$$

[Substituting $\mathbb{P}[\mathcal{M}_{x_0}(x_0) = s]$ from (8)].

B Theoretical outline

In $\mathcal{M}$, we extend the idea of ADP to a non-adapted, general DP by using the highest value of δ across the primary inputs of every member in $\mathfrak{U}$, for a fixed ϵ. This essentially ensures the worst possible tight differential privacy guarantee for the shuffle model. After that, we focus on estimating the original distribution of the primary initial dataset.

Let $\mathcal{R}_{\text{kRR}}^{-1}$ denote the inverse[1] of the probabilistic mechanism $\mathcal{R}_{\text{kRR}}$, which is used as the local randomizer for $\mathcal{M}$. Note that $\mathcal{R}_{\text{kRR}}^{-1}$ and $\mathcal{R}_{\text{kRR}}$ are both $k \times k$ stochastic channels as $|\mathcal{X}| = k$. Staying consistent with our previously developed notations, let us, additionally, introduce $H_{\mathfrak{N}}$ broadcasting the frequencies of the elements in $\mathcal{X}$ after they have been sanitized with $\mathfrak{N}$. In other words, $H_{\mathfrak{N}} = \mathfrak{N}_{\epsilon,\delta}(D_{\mathcal{X}}) = (H_{x_0}, \ldots, H_{x_{k-1}})$, where H_{x_i} is the random variable giving the frequency of x_i after $D_{\mathcal{X}}$ has been obfuscated with $\mathfrak{N}_{\epsilon,\delta}$.

Since both $\mathcal{M}$ and $\mathfrak{N}$ are probabilistic mechanisms, to estimate their utilities we study how accurately we can estimate the true distribution from which $D_{\mathcal{X}}$ is sampled, after observing the response of the histogram queries in both the scenarios.

[1] the inverse of a k-RR mechanism always exists [1,13].

Let $\pi = (\pi_{x_0}, \ldots, \pi_{x_{k-1}})$ be the distribution of the original messages in $D(x_0)$. Our best guess of the original distribution by observing the noisy histogram going through the Gaussian mechanism is the noisy histogram itself, as $\mathbb{E}(H_{x_i}) = n\pi_{x_i}$ for every $i \in \{0, \ldots, k-1\}$.

However, in the case where $D(x_0)$ is locally obfuscated using $\mathcal{R}_{\text{kRR}}$ and the frequency of each element is broadcast by the shuffle model $\mathcal{M}$, we can use the matrix inversion method [1,13] to estimate the distribution of the original messages in $D(x_0)$. So $\mathcal{M}(D(x_0))\mathcal{R}_{\text{kRR}}^{-1}$ (referred as *shuffle+INV* in the experiments) should be giving us $\hat{\pi} = (\hat{\pi}_{x_0}, \ldots, \hat{\pi}_{x_{k-1}})$ – the most likely estimate of the distribution of each user's message in $D(x_0)$ sampled from $\mathcal{X}$ – where $\hat{\pi}_{x_i}$ denotes the random variable estimating the normalised frequency of x_i in $D(x_0)$.

$$\mathbb{E}(\hat{\pi}) = \mathbb{E}(\mathcal{M}(D(x_0))\mathcal{R}_{\text{kRR}}^{-1}) = \pi\mathcal{R}_{\text{kRR}}\mathcal{R}_{\text{kRR}}^{-1} = \pi \tag{12}$$

We recall that $\mathcal{M}$ provides tight (ϵ, δ)-ADP for x_0, x_1, where δ is a function of ϵ_0, ϵ, and x_0 – essentially $\mathcal{M}$ privatizes the true query response for x_0 to be identified as that for any $x_1 \neq x_0$. On the other hand, $\mathfrak{N}_{\epsilon,\delta}$ ensures (ϵ, δ)-DP, which essentially means it guarantees (ϵ, δ)-ADP for every $x_i \in \mathcal{X}$. Therefore, in order to facilitate a fair comparison of utility between the central and shuffle models of differential privacy under the same privacy level for the histogram query, we introduce the following concepts:

i) Individual specific utility: Suppose the primary input of u is x_0. *Individual specific utility* refers to measuring the utility for the specific message x_0 in the dataset $D(x_0)$ in a certain privacy mechanism. In particular, the individual specific utility of x_0 in $D(x_0)$ for $\mathcal{M}$ is

$$\overline{\mathcal{W}}(\mathcal{M}, x_0) = |n\hat{\pi}_{x_0} - n\pi_{x_0}|,$$

and that for $\mathfrak{N}_{\epsilon,\delta}$ is

$$\overline{\mathcal{W}}(\mathfrak{N}_{\epsilon,\delta}, x_0) = |n\pi_{x_0} - H_{x_0}|$$

ii) Community level utility: Here we consider the utility privacy mechanisms over the entire community, i.e., all the values of the original dataset, by measuring the distance between the estimated original distribution obtained from the observed noisy histogram and the original distribution of the source messages itself.

In particular, fixing any $\epsilon_0 > 0$ and $\epsilon > 0$, the *community level utility* for $\mathcal{M}$ is

$$\mathcal{W}(\mathcal{M}) = d(n\hat{\pi}, n\pi), \tag{13}$$

and that for $\mathfrak{N}_{\epsilon,\delta}$[2] is

$$\mathcal{W}(\mathfrak{N}_{\epsilon,\delta}) = d(H_{\mathfrak{N}_{\epsilon,\delta}}, n\pi), \tag{14}$$

where $d(.)$ is any standard metric[3] to measure probability distributions over a finite space.

[2] where δ is correspondingly obtained using Result 1.

[3] we consider Total Variation Distance for our experiments.

For an equitable comparison between $\mathcal{M}$ and $\mathfrak{N}$, we take the worst tight ADP guarantee over every user's primary input and call this the *community level tight DP guarantee for* $\mathcal{M}$. That is, for a fixed ϵ_0, $\epsilon > 0$, we have $\mathcal{M}$ satisfying $(\epsilon, \hat{\delta})$-DP as the community level tight DP guarantee if

$$\hat{\delta} = \max_{x \in \mathcal{X}} \{\delta : \mathcal{M} \text{ is tightly } (\epsilon, \delta(x))\text{-ADP for } x \in D_{\mathcal{X}}\} \tag{15}$$

Therefore, we impose the worst tight ADP guarantee on $\mathcal{M}$ over all the original messages with ϵ and $\hat{\delta}$, implying that $\mathcal{M}$ now gives a $(\epsilon, \hat{\delta})$-DP guarantee by Remark 1, placing us in a position to compare the community level utilities of the shuffle and the central models of DP under the histogram query for a fixed level of privacy. In particular, we juxtapose $\mathcal{W}(\mathcal{M})$ with $\mathcal{W}(\mathfrak{N}_{\epsilon,\hat{\delta}})$, as seen in the experimental results with location data from San Francisco and Paris in Fig. 3.

References

1. Agrawal, R., Srikant, R., Thomas, D.: Privacy preserving olap. In: Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data, pp. 251–262 (2005)
2. Balcer, V., Cheu, A.: Separating local & shuffled differential privacy via histograms. arXiv preprint arXiv:1911.06879 (2019)
3. Balle, B., Bell, J., Gascón, A., Nissim, K.: The privacy blanket of the shuffle model. In: Boldyreva, A., Micciancio, D. (eds.) CRYPTO 2019. LNCS, vol. 11693, pp. 638–667. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26951-7_22
4. Balle, B., Wang, Y.X.: Improving the gaussian mechanism for differential privacy: analytical calibration and optimal denoising. In: International Conference on Machine Learning, pp. 394–403. PMLR (2018)
5. Bittau, A., et al.: Prochlo: strong privacy for analytics in the crowd. In: Proceedings of the 26th Symposium on Operating Systems Principles, pp. 441–459 (2017)
6. Cheu, A.: Differential privacy in the shuffle model: a survey of separations. arXiv preprint arXiv:2107.11839 (2021)
7. Cheu, A., Smith, A., Ullman, J., Zeber, D., Zhilyaev, M.: Distributed differential privacy via shuffling. In: Ishai, Y., Rijmen, V. (eds.) EUROCRYPT 2019. LNCS, vol. 11476, pp. 375–403. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-17653-2_13
8. Duchi, J.C., Jordan, M.I., Wainwright, M.J.: Local privacy and statistical minimax rates. In: 2013 IEEE 54th Annual Symposium on Foundations of Computer Science, pp. 429–438. IEEE (2013)
9. Dwork, C., McSherry, F., Nissim, K., Smith, A.: Calibrating noise to sensitivity in private data analysis. In: Halevi, S., Rabin, T. (eds.) TCC 2006. LNCS, vol. 3876, pp. 265–284. Springer, Heidelberg (2006). https://doi.org/10.1007/11681878_14
10. Erlingsson, Ú., Feldman, V., Mironov, I., Raghunathan, A., Talwar, K., Thakurta, A.: Amplification by shuffling: From local to central differential privacy via anonymity. In: Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 2468–2479. SIAM (2019)
11. Feldman, V., McMillan, A., Talwar, K.: Hiding among the clones: A simple and nearly optimal analysis of privacy amplification by shuffling. arXiv preprint arXiv:2012.12803 (2020)

12. The gowalla dataset. [online]. https://snap.stanford.edu/data/loc-gowalla.html (2011), (Accessed 10 Aug 2021)
13. Kairouz, P., Bonawitz, K., Ramage, D.: Discrete distribution estimation under local privacy. In: International Conference on Machine Learning, pp. 2436–2444. PMLR (2016)
14. Koskela, A., Heikkilä, M.A., Honkela, A.: Tight accounting in the shuffle model of differential privacy. arXiv preprint arXiv:2106.00477 (2021)
15. Sommer, D.M., Meiser, S., Mohammadi, E.: Privacy loss classes: the central limit theorem in differential privacy. Proc. Priv. Enhancing Technol. **2019**(2), 245–269 (2019)