Discrete Applied Mathematics 211 (2016) 86–98

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Differentially private response mechanisms on categorical data

Naoise Holohan [a], Douglas J. Leith [a], Oliver Mason [b,c,*]

[a] School of Computer Science and Statistics, Trinity College Dublin, Ireland

[b] Department of Mathematics and Statistics/Hamilton Institute, Maynooth University-National University of Ireland Maynooth, Co. Kildare, Ireland

[c] Lero, The Irish Software Research Centre, Ireland

ARTICLE INFO

Article history:
Received 28 April 2015
Received in revised form 11 March 2016
Accepted 16 April 2016
Available online 20 May 2016

Keywords:
Data privacy
Differential privacy
Optimal mechanisms

ABSTRACT

We study mechanisms for differential privacy on finite datasets. By deriving *sufficient sets* for differential privacy we obtain necessary and sufficient conditions for differential privacy, a tight lower bound on the maximal expected error of a discrete mechanism and a characterisation of the optimal mechanism which minimises the maximal expected error within the class of mechanisms considered.

1. Introduction

Data privacy has been of interest to researchers for decades [4], but high-profile privacy breaches in recent years, such as those involving AOL [1] and Netflix [18], have renewed focus on the topic. The movement towards smart metering systems for electricity, water and other utilities and the greater use of data mining in so-called smart cities and transport have given rise to further concerns over personal data privacy.

The most traditional framework for the study of data privacy is that of tabular data. A simple model of this type considers the data to be arranged as individual records within a table, where each record contains entries from some underlying dataset, which may be continuous or discrete depending on the type of data being studied. Simple anonymisation techniques, such as removing names and social security numbers (so-called unique identifiers) from the data, have been shown to be inadequate [19]. More sophisticated frameworks such as k-anonymity [19] and ℓ-diversity [15] are also vulnerable to privacy attacks via the use of appropriate side-information or data from external sources [15,14].

Within the last decade, differential privacy [5] has emerged as a popular framework for research in the field of data privacy based on its capability to provide a quantifiable basis for privacy preserving data publishing and mining. This is a probabilistic approach to data privacy in which a suitably *randomised* version of the correct response to a query is released. The core idea is founded on the simple premise that the response to a user query should not be too tightly coupled with any one entry in the table. One widely-adopted implementation of differential privacy for real-valued databases is to add an appropriate amount of noise sampled from a Laplace distribution to each cell of the database [6].

Much research on differential privacy to date has been completed on real-valued databases [6], although a considerable body of literature also exists on discrete data [16,3]; in particular some recent work has focused on graph data relevant to applications in areas such as social networks [13,2].

* Corresponding author. Tel.: +353 01 7083672; fax: +353 501 7083913.
E-mail address: oliver.mason@nuim.ie (O. Mason).

http://dx.doi.org/10.1016/j.dam.2016.04.010
0166-218X/

Differentially private mechanisms can be divided into two distinct classes: sanitisation based mechanisms; and output perturbation based mechanisms. Our concern here is with the former class, which first constructs a sanitised version of the database and then answers queries on this. It has been shown in [11] that if the sanitised database satisfies the requirements of differential privacy, then any query can be answered on it in a differentially private manner.

In writing this paper, we have two aims: the first is to present a set of new results on the mathematical foundations of differential privacy for discrete data; the second is to bring the problems in this field to the attention of researchers in discrete applied mathematics.

We examine differentially private mechanisms for discrete data within the general probabilistic framework described in our previous paper [11]. As we deal with finite datasets here, many of the measure-theoretic details required for the more general setting can be suppressed. However, to properly set context, we include the more general definitions here in Section 2.

Our first results concern an adaptation for discrete data of the exponential mechanism introduced by McSherry and Talwar. In particular, we consider the problem of *sufficient sets* for differential privacy for this mechanism. This problem is motivated by the practical issue of testing whether or not a mechanism is differentially private and arises from the following simple considerations.

For a sanitisation to be differentially private, certain inequalities (described formally later) must hold on all subsets of the database space, which can necessitate checking a prohibitively large collection of sets in order to test for differential privacy. The question of sufficient sets asks whether it is sufficient for the differential privacy condition to hold on a collection of these subsets for it to hold on all subsets. We can therefore reduce the workload required to check that a mechanism satisfies differential privacy. In Section 3, we present results characterising sufficient sets for the discrete exponential mechanism. We then use these to give necessary and sufficient conditions for differential privacy for this mechanism.

A major concern of privacy research is the trade-off between privacy and accuracy. For the current setting, in the absence of a given metric on the dataset, we measure the error of a sanitisation using hamming distance; in Theorem 5 we derive a tight lower bound on the maximal expected error of a discrete exponential mechanism.

In Section 4 we consider a seemingly unrelated approach to database sanitisation: product sanitisations. We show that these are in fact equivalent to the discrete exponential mechanism constructed using the hamming distance and, building on results in [11], we characterise differential privacy and the error for these in Theorems 8 and 9 respectively. Finally in Theorem 10 we provide a characterisation of the optimal product sanitisation mechanism, which minimises the maximal expected error within the class of product sanitisations (and hence within the class of discrete exponential mechanisms). Concluding remarks are given in Section 5.

1.1. Related work

Before the advent of differential privacy, Fienberg examined the use of data swapping and cell suppression for privacy protection on categorical data [9]. Dwork then presented the notion of differential privacy in [5], and its limitations were discussed by Dankar in [7], including its applicability to categorical data.

Dwork's work was closely followed by McSherry and Talwar who proposed the exponential mechanism in [16]. An instantiation of this was used by Hardt and Talwar [10] in examining the geometry of differential privacy. Mohammed made use of the exponential mechanism for releasing count queries in [17,3], while a more recent contribution has looked at differential privacy on counts using a combination of the Laplace and exponential mechanisms [21].

2. Preliminaries

2.1. Database model

We consider a finite data set D with $(m+1)$ elements ($m \geq 1$). A database $\mathbf{d}$ with n rows drawn from this data set is represented by a vector $\mathbf{d} = (d_1, \ldots, d_n) \in D^n$. D is equipped with a σ-algebra, in this case the power set 2^D and D^n inherits the product σ-algebra, 2^{D^n}. We are therefore considering all subsets of D and D^n.

We will consider hamming distance on D^n. Recall that the hamming distance, $\mathrm{h} : D^n \times D^n \to \{0, 1, \ldots, n\}$, between two databases is the number of rows on which they differ:

$$\mathrm{h}(\mathbf{d}, \mathbf{d}') = |\{i : d_i \neq d'_i\}|. \tag{1}$$

Definition 1 (*Neighbours*)**.** Two databases $\mathbf{d}, \mathbf{d}' \in D^n$ are said to be *neighbours*, written $\mathbf{d} \sim \mathbf{d}'$ if $\mathrm{h}(\mathbf{d}, \mathbf{d}') = 1$.

Informally, two databases are neighbours if they differ on exactly one row.

2.2. Query model

We make use of the generalised query model introduced in [11], adapted to the discrete setting. A query $Q : D^n \to E_Q$ outputs a response in E_Q, the structure of which is not specified (it may be numeric, categorical, functional, etc.). E_Q is, however, equipped with a σ-algebra $\mathscr{A}_Q$. We require that all queries be measurable, which is trivial in this setting since $Q^{-1}(A) \subseteq D^n$ for all $A \in \mathscr{A}_Q$.

2.3. Response mechanism

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We define a response mechanism for a query Q to be the family of measurable mappings

$$\{X_{Q,\mathbf{d}} : \Omega \to E_Q \mid \mathbf{d} \in D^n\}. \tag{2}$$

For simplicity, when the query in question is the identity query I, we denote $X_{I,\mathbf{d}}$ by $X_{\mathbf{d}} : \Omega \to D^n$. Where there is no ambiguity, the response mechanism will be written as $\{X_{Q,\mathbf{d}}\}$ (or $\{X_{\mathbf{d}}\}$ when dealing with the identity query).

In this paper we deal exclusively with *sanitised response mechanisms*, where $X_{Q,\mathbf{d}} = Q \circ X_{\mathbf{d}}$. In this case the mechanism is generated by first sanitising the database $\mathbf{d}$ and then answering queries on the sanitised database $X_{\mathbf{d}}$ without any further modification to the data. Hence, a sanitised response mechanism for a query Q is the family of mappings

$$\{X_{Q,\mathbf{d}} = Q \circ X_{\mathbf{d}} : \Omega \to E_Q \mid \mathbf{d} \in D^n\}. \tag{3}$$

One mechanism which we will make use of in this paper is the exponential mechanism, as described by McSherry and Talwar [16]. We next recall the definition of this mechanism for general output spaces; later we will specialise to the case of discrete categorical data.

Definition 2 (*Exponential Mechanism*)**.** Given a query Q, a query output space E_Q, a utility function (which measures the utility of all possible query answers to the database being queried) $u : D^n \times E_Q \to \mathbb{R}$, a measure μ on E_Q and a normalisation constant $C_{\mathbf{d}}$, the exponential mechanism is defined to be the family of mappings $\{X_{Q,\mathbf{d}} : \Omega \to E_Q \mid \mathbf{d} \in D^n\}$, with probability density function with respect to μ given by $C_{\mathbf{d}}^{-1} e^{\epsilon u(\mathbf{d},q)}$ for each $\mathbf{d} \in D^n$ and for $q \in E_Q$.

Comment: McSherry and Talwar refer to the measure μ in the above definition as a *base measure* on the output space. As mentioned in [16], it is often the uniform probability measure on E_Q and is adjusted by the exponential factor $C_{\mathbf{d}}^{-1} e^{\epsilon u(\mathbf{d},q)}$ to reflect the utility of various pairs of values from D^n and E_Q. The measure can potentially be viewed as describing the underlying distribution of values in E_Q. If E_Q is discrete, as in all cases considered here, we can specify μ by $\{\mu(q) : q \in E_Q\}$ and the exponential mechanism has probability mass function

$$\mathbb{P}(X_{Q,\mathbf{d}} = q) = C_{\mathbf{d}}^{-1} e^{\epsilon u(\mathbf{d},q)} \mu(q), \tag{4}$$

where $q \in E_Q$. In fact, we will focus on the situation where Q is the identity map on a finite D^n and μ is given by the uniform measure.

2.4. Differential privacy

We now define what it means for a response mechanism in our framework to be differentially private.

Definition 3 (*Differential Privacy with Respect to a Query*)**.** Let a query $Q : D^n \to E_Q$ and parameters $\epsilon \geq 0$ and $0 \leq \delta \leq 1$ be given. A response mechanism $\{X_{Q,\mathbf{d}} : \Omega \to E_Q \mid \mathbf{d} \in D^n\}$ is (ϵ, δ)-differentially private if

$$\mathbb{P}(X_{Q,\mathbf{d}} \in A) \leq e^{\epsilon} \mathbb{P}(X_{Q,\mathbf{d}'} \in A) + \delta, \tag{5}$$

for all $\mathbf{d} \sim \mathbf{d}' \in D^n$ and for all measurable $A \subseteq E_Q$.

Note that the relation $\mathbf{d} \sim \mathbf{d}'$ is symmetric, so (5) must hold when $\mathbf{d}$ and $\mathbf{d}'$ are swapped.

Note: McSherry and Talwar showed that the exponential mechanism (Definition 2) satisfies $2\epsilon\Delta u$-differential privacy ($\delta = 0$), where

$$\Delta u = \max_{\mathbf{d} \sim \mathbf{d}' \in D^n, q \in E_Q} |u(\mathbf{d}, q) - u(\mathbf{d}', q)|.$$

Theorem 4 of [11] simplifies the problem of checking differential privacy for sanitised response mechanisms. We now recall that theorem.

Theorem 1 (*Identity Query*)**.** *A sanitised response mechanism which is (ϵ, δ)-differentially private with respect to the identity query is (ϵ, δ)-differentially private with respect to any query Q.*

We therefore need only examine the response mechanism for the identity query,

$$\{X_{\mathbf{d}} : \Omega \to D^n \mid \mathbf{d} \in D^n\}. \tag{6}$$

Example 1 (*Categorical Data I*)**.** Suppose the data we are interested in records individuals' favourite hobby. The data set D would contain a list of all hobbies. For simplicity in this example, we restrict answers to the following five hobbies: Sports; Cars; Television; Computer games; and Reading. Hence, $m = 4$. Each database $\mathbf{d}$ would contain the favourite hobby of n individuals. If $n = 6$, one possible $\mathbf{d}$ could be represented by the following list: Sports; Computer games; Television; Sports; Reading; Television.

Queries on such databases could include counting the number of unique hobbies (4 in the case above) or how many list 'Television' as their favourite hobby (2 in the case above). The identity query would be another valid query.

3. Sufficient sets for discrete exponential mechanism

In order for a response mechanism to be deemed (ϵ, δ)-differentially private, (5) must hold for all pairs of neighbouring databases and for all possible subsets of D^n. If we were to check all combinations, this would require checking all $nm(m+1)^n$ pairs of neighbouring databases on $2^{(m+1)^n}-2$ subsets of D^n (all subsets except D^n itself and $\emptyset$). Therefore, checking a discrete response mechanism for differential privacy requires $nm(m+1)^n(2^{(m+1)^n}-2)$ checks in total.

However, for a given pair of neighbouring databases, it is not always necessary to check (5) on all subsets of D^n. So we ask the question: For a given $\mathbf{d} \sim \mathbf{d}'$, what is the smallest collection of subsets of D^n that we need to check for (5) to hold on all subsets of D^n? We call such a collection of subsets *sufficient sets* of the mechanism for $\mathbf{d} \sim \mathbf{d}'$.

In this section, we examine the sufficient sets for a class of discrete response mechanisms and show that, even in the most general cases, significant improvements on workload can be made when checking for differential privacy. We also present conditions that are necessary and sufficient for differential privacy to hold. This compares to the mostly sufficient conditions presented in other differential privacy literature, which can therefore give a conservative estimate on the privacy level achieved.

3.1. General response mechanism

We begin by considering the exponential mechanism described by McSherry and Talwar [16], as detailed in Definition 2. We wish to assign a probability to each database based on its utility to the input/reference database. This is determined by the utility function, which can be a metric or any other function deemed suitable for a particular application. As discussed in Section 2.4, we are only concerning ourselves with the identity query.

With the aim of minimising error, a reasonable approach is to assign the reference/input database itself the highest probability of being returned, with decreasing likelihood the further we move away from the reference as determined by the utility function.

Definition 4 (*Discrete Exponential Mechanism*)**.** Let $u : D^n \times D^n \to \mathbb{R}$ be given. The *discrete exponential response mechanism* is defined to be a family of measurable mappings

$$\{X_{\mathbf{d}} : \Omega \to D^n \mid \mathbf{d} \in D^n\},$$

where each $X_{\mathbf{d}}$ satisfies

$$\mathbb{P}(X_{\mathbf{d}} = \mathbf{d}') = C_{\mathbf{d}}^{-1} e^{u(\mathbf{d},\mathbf{d}')}, \tag{7}$$

for all $\mathbf{d}, \mathbf{d}' \in D^n$. As D^n is finite, we can define the normalisation constant $C_{\mathbf{d}}$ as

$$C_{\mathbf{d}} = \sum_{\mathbf{d}' \in D^n} e^{u(\mathbf{d},\mathbf{d}')}$$

for each $\mathbf{d} \in D^n$.

Remark. While similar to the exponential mechanism (4), the discrete exponential mechanism differs by dealing only with the identity query $(Q = I)$, by having a uniform measure $(\mu = 1)$ and by absorbing ϵ into u, in order to allow consideration of (ϵ, δ)-differential privacy.

Even with this general set-up, we can still make improvements on workload when checking for differential privacy. We begin by defining the following set for each pair of neighbouring databases $\mathbf{d} \sim \mathbf{d}' \in D^n$:

$$\begin{aligned} S_{\mathbf{d},\mathbf{d}'} &= \left\{\mathbf{d}^* \in D^n \mid \mathbb{P}(X_{\mathbf{d}} = \mathbf{d}^*) > \mathbb{P}(X_{\mathbf{d}'} = \mathbf{d}^*)\right\} \\ &= \left\{\mathbf{d}^* \in D^n \mid C_{\mathbf{d}}^{-1} e^{u(\mathbf{d},\mathbf{d}^*)} > C_{\mathbf{d}'}^{-1} e^{u(\mathbf{d}',\mathbf{d}^*)}\right\}. \end{aligned} \tag{8}$$

This set is a collection of the "worst-case" databases for $\mathbf{d}$ and $\mathbf{d}'$, and, as we show in Theorem 2, is the only set of interest when checking for differential privacy.

Theorem 2 (*Sufficient Sets*)**.** *Let $\{X_{\mathbf{d}}\}$ be a discrete exponential mechanism and fix $\mathbf{d} \sim \mathbf{d}' \in D^n$. If (5) holds on all $A \subseteq S_{\mathbf{d},\mathbf{d}'}$ then it will hold on all $A \subseteq D^n$.*

Proof. We fix $\mathbf{d} \sim \mathbf{d}' \in D^n$, let $A \subseteq D^n$ be given and assume (5) holds on all subsets of $S_{\mathbf{d},\mathbf{d}'}$. By assumption, (5) holds on $A_0 = A \cap S_{\mathbf{d},\mathbf{d}'}$, hence $\mathbb{P}(X_{\mathbf{d}} \in A_0) \leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A_0) + \delta$. If $A_0 = A$ we are done, so assume $A \setminus A_0 \neq \emptyset$.

For each $\mathbf{d}^* \in A \setminus A_0$, $\mathbb{P}(X_{\mathbf{d}} = \mathbf{d}^*) \leq \mathbb{P}(X_{\mathbf{d}'} = \mathbf{d}^*)$. Pick one such $\mathbf{d}_0^* \in A \setminus A_0$, then

$$\begin{aligned} \mathbb{P}(X_{\mathbf{d}} \in A_0 \cup \{\mathbf{d}_0^*\}) &= \mathbb{P}(X_{\mathbf{d}} \in A_0) + \mathbb{P}(X_{\mathbf{d}} = \mathbf{d}_0^*) \\ &\leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A_0) + \delta + \mathbb{P}(X_{\mathbf{d}} = \mathbf{d}_0^*) \\ &\leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A_0) + \delta + \mathbb{P}(X_{\mathbf{d}'} = \mathbf{d}_0^*) \\ &\leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A_0 \cup \{\mathbf{d}_0^*\}) + \delta. \end{aligned}$$

Hence, (5) holds on $A_1 = A_0 \cup \{\mathbf{d}_0^*\}$. We can similarly show that (5) holds on $A_2 = A_1 \cup \{\mathbf{d}_1^*\}$ for any $\mathbf{d}_1^* \in A \setminus A_1$. By repeating this process (picking $\mathbf{d}_i^* \in A \setminus A_i$), we can show that (5) holds on $A_{i+1} = A_i \cup \{\mathbf{d}_i^*\}$ for each i.

Since A is finite (D^n is finite), this process will eventually terminate when $A_i = A$, *i.e.* $i = |A \setminus A_0|$. Hence, (5) will hold on A as required. □

We now look at a simple example to demonstrate the impact of Theorem 2.

Example 2 (ℓ^1 *norm*)**.** For this example, we consider a discrete exponential mechanism where $D = \{0, 1, 2\}$, $n = 2$ and $u(\mathbf{d}, \mathbf{d}') = -\|\mathbf{d} - \mathbf{d}'\|_1$. In this case, $|D^n| = 9$, and we are therefore required to check $2^9 - 2 = 510$ subsets (all subsets of D^n except $\emptyset$ and D^n itself) for every pair of neighbouring databases $\mathbf{d} \sim \mathbf{d}' \in D^n$, meaning a total of $510 \times 36 = 18,360$ checks.

Let $\mathbf{d} = \binom{0}{1}$ and $\mathbf{d}' = \binom{2}{1}$, then $S_{\mathbf{d},\mathbf{d}'} = \left\{\binom{0}{0}, \binom{0}{1}, \binom{0}{2}\right\}$. Hence, for this particular pair of neighbouring databases, it is sufficient to check that (5) holds on just $2^3 - 1 = 7$ subsets (all subsets of $S_{\mathbf{d},\mathbf{d}'}$ except $\emptyset$).

If we choose $\mathbf{d} = \binom{1}{1}$ and $\mathbf{d}' = \binom{2}{1}$, then we get $S_{\mathbf{d},\mathbf{d}'} = \left\{\binom{0}{0}, \binom{0}{1}, \binom{0}{2}, \binom{1}{0}, \binom{1}{1}, \binom{1}{2}\right\}$. This gives a total of $2^6 - 1 = 63$ subsets to check.

By populating the entire set of databases, we can show that 21 pairs of neighbour databases require 7 subset checks, while the remaining 15 pairs require 63 checks. That leaves us with a total of 1092 subset checks to verify differential privacy, compared with 18,360 without the use of Theorem 2.

3.2. Response mechanism with fixed $C_{\mathbf{d}}$

By Theorem 2, we know that, for a discrete exponential mechanism, checking that (5) holds on all subsets of $S_{\mathbf{d},\mathbf{d}'}$ is equivalent to checking all subsets of D^n. However, if $C_{\mathbf{d}}$ is fixed for all $\mathbf{d} \in D^n$, we can partition $S_{\mathbf{d},\mathbf{d}'}$ to reduce our workload further. Let us first consider the following set relating to $S_{\mathbf{d},\mathbf{d}'}$:

$$\alpha = \{u(\mathbf{d}, \mathbf{d}^*) - u(\mathbf{d}', \mathbf{d}^*) \mid \mathbf{d}^* \in S_{\mathbf{d},\mathbf{d}'}\}. \tag{9}$$

There are only finitely many such values in α, since $S_{\mathbf{d},\mathbf{d}'}$ is finite. We label these elements $\alpha_1, \alpha_2, \ldots, \alpha_s$, with $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_s$.

We then partition $S_{\mathbf{d},\mathbf{d}'}$ into the collection of subsets $\{\tilde{S}^1_{\mathbf{d},\mathbf{d}'}, \tilde{S}^2_{\mathbf{d},\mathbf{d}'}, \ldots, \tilde{S}^s_{\mathbf{d},\mathbf{d}'}\}$ as follows:

$$\tilde{S}^i_{\mathbf{d},\mathbf{d}'} = \left\{\mathbf{d}^* \in S_{\mathbf{d},\mathbf{d}'} \mid u\left(\mathbf{d}, \mathbf{d}^*\right) - u\left(\mathbf{d}', \mathbf{d}^*\right) = \alpha_i\right\}. \tag{10}$$

Note that for each i and for all $\mathbf{d}^* \in \tilde{S}^i_{\mathbf{d},\mathbf{d}'}$,

$$\begin{aligned}\mathbb{P}(X_{\mathbf{d}'} = \mathbf{d}^*) &= Ce^{u(\mathbf{d}',\mathbf{d}^*)} \\ &= Ce^{u(\mathbf{d},\mathbf{d}^*)-\alpha_i} \\ &= e^{-\alpha_i}\mathbb{P}(X_{\mathbf{d}} = \mathbf{d}^*).\end{aligned} \tag{11}$$

We can now show that if (5) holds on these partitions, it will hold on all subsets of each partition, i.e. the partitions are sufficient sets of themselves.

Theorem 3. *Fix $\mathbf{d} \sim \mathbf{d}' \in D^n$ and let $\{X_{\mathbf{d}}\}$ be a discrete exponential mechanism with $C_{\mathbf{d}} = C$ for all $\mathbf{d} \in D^n$. If (5) holds on $\tilde{S}^i_{\mathbf{d},\mathbf{d}'}$ then it will hold on all $A \subseteq \tilde{S}^i_{\mathbf{d},\mathbf{d}'}$ for each i.*

Proof. Fix $\mathbf{d} \sim \mathbf{d}' \in D^n$. We assume

$$\mathbb{P}(X_{\mathbf{d}} \in \tilde{S}^i_{\mathbf{d},\mathbf{d}'}) \leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in \tilde{S}^i_{\mathbf{d},\mathbf{d}'}) + \delta,$$

and let $A \subseteq \tilde{S}^i_{\mathbf{d},\mathbf{d}'}$. By (11), $\mathbb{P}(X_{\mathbf{d}'} \in A) = e^{-\alpha_i}\mathbb{P}(X_{\mathbf{d}} \in A)$. Since this also holds for the set $\tilde{S}^i_{\mathbf{d},\mathbf{d}'}$ itself, we have

$$1 \leq e^{\epsilon-\alpha_i} + \frac{\delta}{\mathbb{P}(X_{\mathbf{d}} \in \tilde{S}^i_{\mathbf{d},\mathbf{d}'})}.$$

Clearly $\mathbb{P}(X_{\mathbf{d}} \in A) \leq \mathbb{P}(X_{\mathbf{d}} \in \tilde{S}^i_{\mathbf{d},\mathbf{d}'})$, which gives

$$1 \leq e^{\epsilon-\alpha_i} + \frac{\delta}{\mathbb{P}(X_{\mathbf{d}} \in A)},$$

or, rewriting,

$$\mathbb{P}(X_{\mathbf{d}} \in A) \leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A) + \delta.$$

Hence, (5) holds on all $A \subseteq \tilde{S}^i_{\mathbf{d},\mathbf{d}'}$. □

When $\delta = 0$, Theorem 3 tells us that these partitions $\tilde{S}^i_{\mathbf{d},\mathbf{d}'}$ of $S_{\mathbf{d},\mathbf{d}'}$ are all we need to check to verify ϵ-differential privacy, i.e. the sufficient sets are the collection of subsets $\{\tilde{S}^1_{\mathbf{d},\mathbf{d}'}, \ldots, \tilde{S}^s_{\mathbf{d},\mathbf{d}'}\}$.

Corollary 1 (*Sufficient Sets with Fixed $C_{\mathbf{d}}$*)**.** *Let $\{X_{\mathbf{d}}\}$ be a discrete exponential mechanism, $\delta = 0$ and fix $\mathbf{d} \sim \mathbf{d}' \in D^n$. If (5) holds on $\tilde{S}^i_{\mathbf{d},\mathbf{d}'}$ for all i, then it will hold on all subsets of D^n.*

Proof. Let $A \subseteq S_{\mathbf{d},\mathbf{d}'}$ and assume that (5) holds on $\tilde{S}^i_{\mathbf{d},\mathbf{d}'}$ for all i. Hence, by Theorem 3, (5) holds on all subsets of $\tilde{S}^i_{\mathbf{d},\mathbf{d}'}$ for all i. As $\{\tilde{S}^i_{\mathbf{d},\mathbf{d}'}\}$ partitions $S_{\mathbf{d},\mathbf{d}'}$, $A = \bigcup_i A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'}$, hence $\mathbb{P}(X_{\mathbf{d}} \in A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'}) \leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'})$ for all i.

Since $\tilde{S}^i_{\mathbf{d},\mathbf{d}'} \cap \tilde{S}^j_{\mathbf{d},\mathbf{d}'} = \emptyset$ when $i \neq j$, $\mathbb{P}(X_{\mathbf{d}} \in \bigcup_i A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'}) = \sum_i \mathbb{P}(X_{\mathbf{d}} \in A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'})$, and so,

$$\begin{aligned}
\mathbb{P}(X_{\mathbf{d}} \in A) &= \mathbb{P}\left(X_{\mathbf{d}} \in \bigcup_i A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'}\right) \\
&= \sum_i \mathbb{P}\left(X_{\mathbf{d}} \in A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'}\right) \\
&\leq e^{\epsilon} \sum_i \mathbb{P}\left(X_{\mathbf{d}'} \in A \cap \tilde{S}^i_{\mathbf{d},\mathbf{d}'}\right) \\
&= e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A).
\end{aligned}$$

Therefore (5) holds on all subsets of $S_{\mathbf{d},\mathbf{d}'}$, and by Theorem 2, it holds on all subsets of D^n. □

3.3. Discrete exponential mechanism with hamming distance

The usefulness of Theorem 3 becomes particularly apparent when we restrict our response mechanism to one derived from hamming distance. For the remainder of this section our utility function u is defined to be

$$u(\mathbf{d}, \mathbf{d}') = -k\,\mathrm{h}(\mathbf{d}, \mathbf{d}'), \tag{12}$$

where $k \geq 0$ is a privacy parameter and $\mathrm{h}(\mathbf{d}, \mathbf{d}')$ is the hamming distance between $\mathbf{d}$ and $\mathbf{d}'$ (the number of elements on which they differ, see (1)). Note that for this set-up, the normalisation constant $C_{\mathbf{d}} = C$ is fixed for all $\mathbf{d} \in D^n$.

Proposition 1. *For a discrete exponential mechanism satisfying (12),*

$$C_{\mathbf{d}} = C = \left(1 + \frac{m}{e^k}\right)^{-n}. \tag{13}$$

Proof. Let $\mathbf{d} \in D^n$ be given. Then,

$$\begin{aligned}
\sum_{\mathbf{d}' \in D^n} \mathbb{P}(X_{\mathbf{d}} = \mathbf{d}') &= \sum_{\mathbf{d}' \in D^n} C_{\mathbf{d}} e^{-k\,\mathrm{h}(\mathbf{d},\mathbf{d}')} \\
&= C_{\mathbf{d}} e^0 + C_{\mathbf{d}} nme^{-k} + C_{\mathbf{d}} \binom{n}{2} m^2 e^{-2k} + \cdots + C_{\mathbf{d}} m^n e^{-nk} \\
&= C_{\mathbf{d}} \sum_{i=0}^{n} \binom{n}{i} m^i e^{-ik} \\
&= C_{\mathbf{d}} \left(1 + \frac{m}{e^k}\right)^n.
\end{aligned}$$

For the probability mass function of $X_{\mathbf{d}}$ to sum to 1, we need

$$C_{\mathbf{d}} \left(1 + \frac{m}{e^k}\right)^n = 1.$$

Rearranging this completes the proof. □

In this case, the sufficient sets for the problem condense down to a single set.

Corollary 2 (*Sufficient Sets for Hamming Distance*)**.** *Consider a discrete exponential mechanism satisfying (12). If (5) holds on $S_{\mathbf{d},\mathbf{d}'}$, then it will hold on all $A \subseteq D^n$.*

Proof. First note that, for all $\mathbf{d} \sim \mathbf{d}' \in D^n$,

$$\mathrm{h}(\mathbf{d}', \mathbf{d}^*) - \mathrm{h}(\mathbf{d}, \mathbf{d}^*) \in \{-1, 0, 1\},$$

hence for all $\mathbf{d}^* \in S_{\mathbf{d},\mathbf{d}'}$, $\mathrm{h}(\mathbf{d}', \mathbf{d}^*) - \mathrm{h}(\mathbf{d}, \mathbf{d}^*) = 1$. The set α reduces to a singleton set,

$$\begin{aligned}\alpha &= \{-k\,\mathrm{h}(\mathbf{d}, \mathbf{d}^*) + k\,\mathrm{h}(\mathbf{d}', \mathbf{d}^*) \mid \mathbf{d}^* \in S_{\mathbf{d},\mathbf{d}'}\} \\ &= \{k\},\end{aligned}$$

and hence $S_{\mathbf{d},\mathbf{d}'} = \tilde{S}^1_{\mathbf{d},\mathbf{d}'}$.

We can then conclude that if (5) holds on $S_{\mathbf{d},\mathbf{d}'}$, it must hold on all subsets of $S_{\mathbf{d},\mathbf{d}'}$ (by Theorem 2) and also on all subsets of D^n (by Theorem 3). □

We have established that, for the discrete exponential mechanism satisfying (12) and for a given neighbouring pair of databases $\mathbf{d} \sim \mathbf{d}'$, to check for differential privacy on all possible subsets of the database space D^n, we need only check a single set $S_{\mathbf{d},\mathbf{d}'}$.

It is now a relatively simple task to establish conditions on the response mechanism for differential privacy. For the following theorem, we assume that $\delta < 1$ (note that all mechanisms are trivially $(\epsilon,1)$-differentially private).

Theorem 4 (*Condition for Differential Privacy*)**.** *Let $\delta < 1$. A discrete exponential mechanism $\{X_{\mathbf{d}}\}$ satisfying (12) is (ϵ, δ)-differentially private if and only if*

$$e^k \leq \frac{e^\epsilon + m\delta}{1-\delta}. \tag{14}$$

Proof. Fix $\mathbf{d} \sim \mathbf{d}' \in D^n$. By Corollary 2 we need only satisfy (5) on $S_{\mathbf{d},\mathbf{d}'}$ for it to hold on all $A \subseteq D^n$. Hence we need

$$\mathbb{P}(X_{\mathbf{d}} \in S_{\mathbf{d},\mathbf{d}'}) \leq e^\epsilon \mathbb{P}(X_{\mathbf{d}'} \in S_{\mathbf{d},\mathbf{d}'}) + \delta.$$

By definition, $\mathbb{P}(X_{\mathbf{d}'} \in S_{\mathbf{d},\mathbf{d}'}) = e^{-k}\mathbb{P}(X_{\mathbf{d}} \in S_{\mathbf{d},\mathbf{d}'})$, therefore the mechanism will be differentially private if and only if

$$1 \leq e^{\epsilon-k} + \frac{\delta}{\mathbb{P}(X_{\mathbf{d}} \in S_{\mathbf{d},\mathbf{d}'})}. \tag{15}$$

Now consider $\mathbb{P}(X_{\mathbf{d}} \in S_{\mathbf{d},\mathbf{d}'})$. By definition, each $\mathbf{d}^* \in S_{\mathbf{d},\mathbf{d}'}$ lies a hamming distance of $\mathrm{h}(\mathbf{d}, \mathbf{d}^*) + 1$ from $\mathbf{d}'$. Therefore, the number of $\mathbf{d}^*$ where $\mathrm{h}(\mathbf{d}, \mathbf{d}^*) = c$ is $\binom{n-1}{c}m^c$ (c elements must be changed, but not the element on which $\mathbf{d}$ and $\mathbf{d}'$ differ, to ensure $\mathrm{h}(\mathbf{d}, \mathbf{d}^*) + 1 = \mathrm{h}(\mathbf{d}', \mathbf{d}^*)$). Hence,

$$\begin{aligned}\mathbb{P}(X_{\mathbf{d}} \in S_{\mathbf{d},\mathbf{d}'}) &= C\sum_{i=0}^{n-1}\binom{n-1}{i}m^i e^{-ik} \\ &= C\left(1+\frac{m}{e^k}\right)^{n-1} \\ &= \left(1+\frac{m}{e^k}\right)^{-1}.\end{aligned}$$

Substituting this result into (15) gives

$$1 \leq e^{\epsilon-k} + \delta\left(1+\frac{m}{e^k}\right),$$

and solving for e^k completes the proof. □

Remark. For $(\epsilon,0)$-differential privacy, we require $k \leq \epsilon$.

Discussion: If we convert our discrete exponential mechanism back into the form of the exponential mechanism, McSherry and Talwar [16] tell us that the mechanism satisfies 2ϵ-differential privacy at worst. However, we have shown in Theorem 4 that the mechanism satisfies ϵ-differential privacy, and that this condition is tight (necessary and sufficient, hence we can do no better). This improves on the looser bound in McSherry and Talwar's proof, as it underestimates the differential privacy achieved by a factor of two. Their mechanism is also limited to ϵ-differential privacy only ($\delta = 0$), whereas we can account for (ϵ, δ)-differential privacy ($\delta > 0$).

Using the differential privacy constraints established in Theorem 4, we can now determine error bounds for the mechanism, where we define error to be the largest mean hamming distance between the input and sanitised databases. Formally, we have the following result.

Theorem 5 (*Error*)**.** *Consider a discrete exponential mechanism satisfying (12) which is (ϵ, δ)-differentially private. Let*

$$\mathcal{E} = \max_{\mathbf{d}\in D^n} \mathbb{E}[\mathrm{h}(X_{\mathbf{d}}, \mathbf{d})]. \tag{16}$$

Then

$$\frac{1-\delta}{1+\frac{e^\epsilon}{m}} \leq \frac{\mathcal{E}}{n} \leq \frac{m}{m+1}. \tag{17}$$

Proof. Let $\mathbf{d} \in D^n$. Then $\mathbb{P}(X_{\mathbf{d}} = \mathbf{d}') = Ce^{-k\,\mathrm{h}(\mathbf{d},\mathbf{d}')}$ and $|\{\mathbf{d}' : \mathrm{h}(\mathbf{d}, \mathbf{d}') = c\}| = \binom{n}{c}m^c$ for all $c \in \{0, 1, \ldots, n\}$. Hence,

$$\begin{aligned}
\mathbb{E}[\mathrm{h}(X_{\mathbf{d}}, \mathbf{d})] &= C\sum_{i=0}^{n} i\binom{n}{i}m^i e^{-ik} \\
&= Cn\sum_{i=1}^{n}\binom{n-1}{i-1}\left(\frac{m}{e^k}\right)^i \\
&= Cn\frac{m}{e^k}\sum_{l=0}^{n-1}\binom{n-1}{l}\left(\frac{m}{e^k}\right)^l \\
&= Cn\frac{m}{e^k}\left(1+\frac{m}{e^k}\right)^{n-1} \\
&= n\frac{m}{e^k}\left(1+\frac{m}{e^k}\right)^{-1} \\
&= \frac{n}{1+\frac{e^k}{m}}.
\end{aligned}$$

On average, the number of entries that will change is therefore

$$\frac{\mathcal{E}}{n} = \frac{\max_{\mathbf{d}\in D^n}\mathbb{E}[\mathrm{h}(X_{\mathbf{d}}, \mathbf{d})]}{n} = \frac{1}{1+\frac{e^k}{m}}.$$

Since the response mechanism is differentially private, Theorem 4 tells us that $e^k \leq \frac{e^\epsilon+m\delta}{1-\delta}$, and $k \geq 0$ by definition, hence,

$$\frac{1-\delta}{1+\frac{e^\epsilon}{m}} \leq \frac{\mathcal{E}}{n} \leq \frac{m}{m+1}. \quad \square$$

Remark. This lower bound is tight and can be achieved by setting $k = \ln\left(\frac{e^\epsilon+m\delta}{1-\delta}\right)$ (see Theorem 4).

4. Product sanitisation

The results of Section 3 give a clear framework on how to create differentially private mechanisms for any type of discrete data, particularly categorical data using hamming distance, and to obtain tight (necessary and sufficient) conditions for differential privacy. However, this mechanism requires the creation of a separate probability distribution for each of the $(m+1)^n$ unique databases.

In this section, we present a simple method for realising the discrete exponential mechanism.

4.1. Response mechanism

A product sanitisation mechanism is one where the database is sanitised row-by-row. The following is the definition of such a mechanism, following the same notation as in [11].

Definition 5 (*Product Sanitisation Mechanism*)**.** Given a 1-dimensional response mechanism for the identity query, $\{X_d : \Omega \to D \mid d \in D\}$ (the parent mechanism), the *product sanitisation response mechanism* is defined to be the set of measurable mappings

$$\{X_{\mathbf{d}} : \Omega \to D^n \mid \mathbf{d} \in D^n\}$$

given by

$$X_{\mathbf{d}} = (X_{\mathbf{d}}^1, \ldots, X_{\mathbf{d}}^n), \tag{18}$$

where the $X_{\mathbf{d}}^i$ are independent and each $X_{\mathbf{d}}^i$ has the same distribution as X_{d_i}, for all $\mathbf{d} \in D^n$, $i \in \{1, \ldots, n\}$.

In this setting, each row of the sanitised database is represented by an independent random variable, as if the database represents n 1-dimensional databases, each sanitised independently. Realising this framework therefore only requires the creation of $m+1$ probability distributions which are then copied to create distributions on each database.

Remark. The above definition of product sanitisation involves generating "noisy" versions of the data contributions or responses from individuals. In this way, it is clearly related to the classical method of *randomised response* in the literature on statistical disclosure control [20]. Recently, differentially private mechanisms based on randomised responses have been considered in [8], motivated by applications to statistical crowd-sourcing. The techniques described in this latter paper have been deployed in Google's open source browser, Chromium. In [12], the authors study product mechanisms under the heading *local differential privacy*. They consider strict differential privacy ($\delta = 0$) and a low privacy regime. For this setting, they show in Theorem 2.6 of [12] that a particular randomised response mechanism is optimal with respect to a utility function based on Kullback–Leibler divergence. We consider related questions here for the case of relaxed differential privacy and objective functions based on Hamming distance.

We now recall Theorem 5 of [11], which states that differential privacy on a product sanitisation mechanism is guaranteed when its parent mechanism is differentially private.

Theorem 6. *Consider a family $\{X_d \mid d \in D\}$ of measurable mappings and assume that*

$$\mathbb{P}(X_d \in A) \leq e^{\epsilon}\mathbb{P}(X_{d'} \in A) + \delta,$$

for all $d, d' \in D$ and all $A \subseteq D$. Let $\{X_{\mathbf{d}} \mid \mathbf{d} \in D^n\}$ be a product sanitisation mechanism. Then

$$\mathbb{P}(X_{\mathbf{d}} \in A) \leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A) + \delta,$$

for all $\mathbf{d} \sim \mathbf{d}' \in D^n$ and all $A \subseteq D^n$.

The converse of this result was proven in Lemma 3 of [11]. However, we are able to take a simpler approach to showing this converse by exploiting the finiteness of D.

Corollary 3 (*Parent Mechanism*)**.** *A product sanitisation mechanism $\{X_{\mathbf{d}} \mid \mathbf{d} \in D^n\}$ is (ϵ, δ)-differentially private if and only if its parent mechanism $\{X_d \mid d \in D\}$ is (ϵ, δ)-differentially private.*

Proof. "$\Leftarrow$": Theorem 6.

"$\Rightarrow$": Let $A \subseteq D$ and $d \neq d' \in D$ be given. Define $A' = A \times D^{n-1} \subseteq D^n$ and $\mathbf{d}, \mathbf{d}' \in D^n$, such that $d_1 = d$, $d'_1 = d'$ and $d_i = d'_i$ for all $i \in \{2, \ldots, n\}$, hence $\mathbf{d} \sim \mathbf{d}'$. Since $\{X_{\mathbf{d}}\}$ is differentially private by assumption,

$$\mathbb{P}(X_{\mathbf{d}} \in A') \leq e^{\epsilon}\mathbb{P}(X_{\mathbf{d}'} \in A') + \delta.$$

However, since $\{X_{\mathbf{d}}\}$ is a product sanitisation mechanism,

$$\begin{aligned}\mathbb{P}(X_{\mathbf{d}} \in A') &= \mathbb{P}(X^1_{\mathbf{d}} \in A) \times \prod_{i=2}^{n} \mathbb{P}(X^i_{\mathbf{d}} \in D)\\ &= \mathbb{P}(X^1_{\mathbf{d}} \in A)\\ &= \mathbb{P}(X_{d_1} \in A).\end{aligned}$$

Hence,

$$\mathbb{P}(X_d \in A) \leq e^{\epsilon}\mathbb{P}(X_{d'} \in A) + \delta,$$

for all $d \neq d' \in D$ and $A \subseteq D$, as required. □

Hence, for a product sanitisation mechanism to be differentially private, we need only show that its parent mechanism is differentially private.

For the remainder of this section, given $0 \leq p \leq \frac{1}{m+1}$, the parent mechanism $\{X_d \mid d \in D\}$ of the product sanitisation is defined such that

$$\mathbb{P}(X_d = d) = 1 - pm, \qquad \mathbb{P}(X_d = d') = p, \tag{19}$$

for every $d' \in D \setminus \{d\}$.

Remark. $p = \frac{1}{m+1}$ represents the case of releasing uniform noise (i.e. no information), while decreasing p reduces the error of the mechanism.

Remark. By requiring $p \leq \frac{1}{m+1}$, we get $\mathbb{P}(X_d = d) \geq \mathbb{P}(X_d = d')$ for every $d' \in D$. Therefore, the mechanism is at least as likely to return the correct answer as any one incorrect answer, an entirely reasonable assumption.

Example 3 (*Categorical Data II*)**.** Using the same set-up as in Example 1, one such permissable p would be $p = 0.1$, since $0 \leq 0.1 \leq \frac{1}{5}$. Then, if d were to be the value 'Television', and d' the value 'Cars', $\mathbb{P}(X_d = d) = 0.6$, and $\mathbb{P}(X_d = d') = 0.1$.

We then sanitise the n-row database $\mathbf{d}$ in the same way, working through the database one row at a time.

In the first main result of this section, we show that the product sanitisation mechanism and the discrete exponential mechanism are equivalent, despite being constructed in different ways (this is subject to the discrete exponential mechanism satisfying (12) and the product sanitisation mechanism satisfying (19)).

Theorem 7 (*Equivalence*). *Let* $\{X_{\mathbf{d}} \mid \mathbf{d} \in D^n\}$ *be a discrete exponential mechanism satisfying* (12)*, and* $\{Y_{\mathbf{d}} \mid \mathbf{d} \in D^n\}$ *be a product sanitisation response mechanism, whose parent mechanism satisfies* (19)*. Then the probability mass functions of* $X_{\mathbf{d}}$ *and* $Y_{\mathbf{d}}$ *are identical, for every* $\mathbf{d} \in D^n$*, when*

$$e^k = \frac{1}{p} - m. \tag{20}$$

Proof. Let $\mathbf{d} \in D^n$, then

$$\mathbb{P}(X_{\mathbf{d}} = \mathbf{d}') = \left(1 + \frac{m}{e^k}\right)^{-n} e^{-k\,\mathrm{h}(\mathbf{d},\mathbf{d}')},$$

$$\mathbb{P}(Y_{\mathbf{d}} = \mathbf{d}') = (1 - pm)^n \left(\frac{p}{1 - pm}\right)^{\mathrm{h}(\mathbf{d},\mathbf{d}')},$$

for all $\mathbf{d}' \in D^n$.

For the two mechanisms to be equivalent, we need $\frac{p}{1-pm} = e^{-k}$, or, rewriting, $\frac{1-pm}{p} = e^k$.

We also need $1 - pm = \frac{e^k}{e^k+m}$, which can be rewritten as $\frac{1}{1-pm} = 1 + \frac{m}{e^k}$, or $\frac{1-pm}{p} = e^k$.

Hence, the mechanisms $\{X_{\mathbf{d}}\}$ and $\{Y_{\mathbf{d}}\}$ are equivalent when $e^k = \frac{1}{p} - m$ or $p = \frac{1}{e^k+m}$. □

4.2. Alternative proofs of Theorems 4 and 5

We have already established that the discrete exponential mechanism satisfying (12) and the product sanitisation mechanism satisfying (19) are equivalent, meaning the results of Theorems 4 and 5 also apply to the product sanitisation mechanism in this particular set-up. In this sub-section, we provide alternative methods of proof for these theorems that make use of the specific product sanitisation of the mechanism, for the additional insight provided.

The proof of the following theorem first appeared in [11], but is included here for completeness.

Theorem 8 (*Condition for Differential Privacy*). *A product sanitisation mechanism, whose parent mechanism satisfies* (19)*, is* (ϵ, δ)*-differentially private if and only if*

$$p \geq \frac{1 - \delta}{e^\epsilon + m}. \tag{21}$$

Proof. By Corollary 3, we need only be concerned with proving differential privacy on $\{X_d \mid d \in D\}$ for it to hold on $\{X_{\mathbf{d}} \mid \mathbf{d} \in D^n\}$.

"⇒": Assume $\{X_d\}$ is differentially private and let $d \neq d' \in D$ be given. Applying the definition of differential privacy to the singleton set $A = \{d\}$ gives

$$1 - pm \leq e^\epsilon p + \delta.$$

Rearranging this gives (21).

"⇐": Assume (21) holds and let $d \neq d' \in D$ and $A \subseteq D$ be given. There are four cases to consider on A:

1. $d, d' \notin A$: Then $\mathbb{P}(X_d \in A) = \mathbb{P}(X_{d'} \in A) = p|A|$ and differential privacy holds trivially.
2. $d, d' \in A$: Then $\mathbb{P}(X_d \in A) = \mathbb{P}(X_{d'} \in A) = p(|A| - 1) + 1 - pm = p(|A| - m - 1) + 1$ and differential privacy holds trivially.
3. $d' \in A, d \notin A$: Then $\mathbb{P}(X_d \in A) = p|A|$ and $\mathbb{P}(X_{d'} \in A) = p(|A| - 1) + 1 - pm$, so $\mathbb{P}(X_d \in A) \leq \mathbb{P}(X_{d'} \in A)$ and differential privacy holds.
4. $d \in A, d' \notin A$: Then

$$\mathbb{P}(X_d \in A) = p(|A| - m - 1) + 1,$$
$$\mathbb{P}(X_{d'} \in A) = p|A|.$$

From (21), we have

$$\begin{aligned} 1 - pm &\leq e^\epsilon p + \delta \\ &= e^\epsilon (p|A| - p|A| + p) + \delta \\ &\leq e^\epsilon p|A| - p(|A| - 1) + \delta, \end{aligned}$$

since $|A| \geq 1$ ($d \in A$ by hypothesis). Rewriting the above,

$$p(|A| - m - 1) + 1 \leq e^{\epsilon}\,(p|A|) + \delta,$$

hence $\mathbb{P}(X_d \in A) \leq e^{\epsilon}\mathbb{P}(X_{d'} \in A) + \delta$ and differential privacy holds. □

Remark. For $(\epsilon,0)$-differential privacy, we require $p \geq \frac{1}{e^{\epsilon}+m}$.

Remark. By definition, p is bounded from above by $\frac{1}{m+1}$, so for differential privacy we require

$$\frac{1-\delta}{e^{\epsilon}+m} \leq p \leq \frac{1}{m+1}.$$

We now look at an alternative proof of Theorem 5 which established error bounds on the mechanism.

Theorem 9 (*Error*). *The error, $\mathcal{E}$ (given by (16)), of a product sanitisation mechanism, whose parent mechanism satisfies (19) and is $(\epsilon,\ \delta)$-differentially private, satisfies*

$$\frac{1-\delta}{1+\frac{e^{\epsilon}}{m}} \leq \frac{\mathcal{E}}{n} \leq \frac{m}{m+1}. \tag{22}$$

Proof. Let $\mathbf{d} \in D^n$. Then,

$$\begin{aligned}
\mathcal{E} &= \max_{\mathbf{d}\in D^n} \mathbb{E}\left[\mathrm{h}(X_{\mathbf{d}}, \mathbf{d})\right] \\
&= \max_{\mathbf{d}\in D^n} \mathbb{E}\left[\sum_{i=1}^{n} \mathrm{h}(X^i_{\mathbf{d}}, d_i)\right] \\
&= \max_{\mathbf{d}\in D^n} \sum_{i=1}^{n} \mathbb{E}\left[\mathrm{h}(X^i_{\mathbf{d}}, d_i)\right] \\
&= \max_{\mathbf{d}\in D^n} \sum_{i=1}^{n} \mathbb{P}(X_{d_i} \neq d_i) \\
&= n\left(\max_{d\in D}\left(1 - \mathbb{P}(X_d = d)\right)\right) \\
&= npm.
\end{aligned}$$

On average, the number of entries that will change is therefore

$$\frac{\mathcal{E}}{n} = pm.$$

As the mechanism satisfies $(\epsilon,\ \delta)$-differential privacy, $p \geq \frac{1-\delta}{e^{\epsilon}+m}$ by Theorem 8, and since $p \leq \frac{1}{m+1}$ by definition,

$$\frac{1-\delta}{1+\frac{e^{\epsilon}}{m}} \leq \frac{\mathcal{E}}{n} \leq \frac{m}{m+1}.$$ □

Remark. As with the discrete exponential mechanism, this bound is tight and can be achieved by setting p equal to the lower bound established in Theorem 8.

4.3. Optimal mechanism

We now look at the second main result of this section. Using the same error definition as before, (16), we show how to construct the optimal $(\epsilon,\ \delta)$-differentially private mechanism which produces the minimum error. For the purpose of this subsection, we assume the following labelling of elements in the data set:

$$D = \{\tilde{d}_1, \tilde{d}_2, \ldots, \tilde{d}_{m+1}\}. \tag{23}$$

Definition 6 (*Solution Matrix*). The parent mechanism $\{X_d\}$ of a product sanitisation mechanism can be defined by a stochastic matrix

$$P_{(\epsilon,\delta)} \in [0,\ 1]^{(m+1)\times(m+1)},$$

where

$$\mathbb{P}(X_{\tilde{d}_i} = \tilde{d}_j) = \left[P_{(\epsilon,\delta)}\right]_{ij}. \tag{24}$$

We refer to $P_{(\epsilon,\delta)}$ as a product sanitisation *solution matrix* if the mechanism it defines is $(\epsilon,\ \delta)$-differentially private.

Remark. A parent mechanism which satisfies (19) is represented by the following solution matrix:

$$\left[P_{(\epsilon,\delta)}\right]_{ij} = \begin{cases} 1 - mp & \text{if } i = j, \\ p & \text{if } i \neq j. \end{cases} \tag{25}$$

Theorem 10 (*Optimality*)**.** *Let $\{X_d\}$ be a parent mechanism which satisfies (19), with $p = \frac{1-\delta}{e^\epsilon+m}$. The error, $\mathcal{E} = \max_{\mathbf{d}\in D^n} \mathbb{E}[h(X_{\mathbf{d}}, \mathbf{d})]$, produced by its product sanitisation mechanism $\{X_{\mathbf{d}}\}$ is the minimum of all product sanitisation mechanisms.*

Proof. Let $A = P_{(\epsilon,\delta)}$, satisfying (25) with $p = \frac{1-\delta}{e^\epsilon+m}$, be the solution matrix of $\{X_d\}$, *i.e.* $\mathbb{P}(X_{\tilde{d}_i} = \tilde{d}_j) = a_{ij}$. Note that this mechanism has the property that $a_{jj} = e^\epsilon a_{ij} + \delta$, since $a_{jj} = \frac{e^\epsilon+m\delta}{e^\epsilon+m}$ and $a_{ij} = \frac{1-\delta}{e^\epsilon+m}$, for all $i \neq j$.

The error of its product sanitisation mechanism $\{X_{\mathbf{d}}\}$, using the same method as in the proof of Theorem 9, is

$$\begin{aligned} \max_{\mathbf{d}\in D^n} \mathbb{E}[\mathrm{h}(X_{\mathbf{d}}, \mathbf{d})] &= n\left(\max_{d\in D}(1 - \mathbb{P}(X_d = d))\right) \\ &= n\left(\max_i (1 - a_{ii})\right) \\ &= n\left(1 - \min_i a_{ii}\right) \\ &= n(1 - a_{ii}), \end{aligned}$$

for all i, since $a_{ii} = a_{jj} = 1 - mp$ for all i and j.

Let B be a solution matrix defining the parent mechanism $\{Y_d\}$, with a corresponding product sanitisation mechanism $\{Y_{\mathbf{d}}\}$, where $B \neq A$. Since B is a solution matrix, it is stochastic ($\sum_j b_{ij} = 1$ for all i) and the parent mechanism it defines is $(\epsilon,\ \delta)$-differentially private ($\mathbb{P}(Y_d \in A) \leq e^\epsilon \mathbb{P}(Y_{d'} \in A) + \delta$, for all $d, d' \in D$ and $A \subseteq D$).

Since $A \neq B$ and A, B are stochastic, there exists at least one pair (i^*, j^*), where $b_{i^*j^*} < a_{i^*j^*}$. There are two cases to consider:

1. $i^* = j^*$: The error of $\{Y_{\mathbf{d}}\}$ is then

$$\begin{aligned} \max_{\mathbf{d}\in D^n} \mathbb{E}[\mathrm{h}(Y_{\mathbf{d}}, \mathbf{d})] &= n\left(1 - \min_i b_{ii}\right) \\ &\geq n\left(1 - b_{j^*j^*}\right) \\ &> n(1 - a_{j^*j^*}) \\ &= \max_{\mathbf{d}\in D^n} \mathbb{E}[\mathrm{h}(X_{\mathbf{d}}, \mathbf{d})]. \end{aligned}$$

2. $i^* \neq j^*$: As noted previously, $a_{j^*j^*} = e^\epsilon a_{i^*j^*} + \delta$, and since $\{Y_d\}$ is $(\epsilon,\ \delta)$-differentially private, $\mathbb{P}(Y_{d_{j^*}} = d_{j^*}) \leq e^\epsilon \mathbb{P}(Y_{d_{i^*}} = d_{j^*}) + \delta$, or alternatively, $b_{j^*j^*} \leq e^\epsilon b_{i^*j^*} + \delta$. Therefore,

$$\begin{aligned} b_{j^*j^*} &\leq e^\epsilon b_{i^*j^*} + \delta \\ &< e^\epsilon a_{i^*j^*} + \delta \\ &= a_{j^*j^*}. \end{aligned}$$

Hence, $b_{j^*j^*} < a_{j^*j^*}$, and case 1 applies.

We therefore conclude that

$$\max_{\mathbf{d}\in D^n} \mathbb{E}[\mathrm{h}(Y_{\mathbf{d}}, \mathbf{d})] > \max_{\mathbf{d}\in D^n} \mathbb{E}[\mathrm{h}(X_{\mathbf{d}}, \mathbf{d})],$$

and that $\{X_{\mathbf{d}}\}$ is the product sanitisation mechanism which produces the optimal error. □

Using Theorem 10, we now have a simple method to construct the most accurate $(\epsilon,\ \delta)$-differentially private mechanism possible for discrete data. We have proven that no other product sanitisation mechanism is more accurate than it. It follows from Theorem 7 that we now know the optimal discrete exponential mechanism that satisfies (12).

Remark. The previous result describes an optimal randomised response mechanism for what is referred to as *local differential privacy* in [12]. Where this earlier paper considered an objective function based on Kullback–Leibler divergence and the case $\delta = 0$, our result addresses the relaxed formulation of differential privacy with $\delta \geq 0$ and an objective function based on Hamming distance.

5. Conclusion

We study mechanisms for (ϵ, δ)-differential privacy on finite datasets that are an adaptation for discrete data of the exponential mechanism introduced by McSherry and Talwar. By deriving *sufficient sets* for differential privacy we obtain necessary and sufficient conditions for differential privacy, a tight lower bound on the maximal expected error of a discrete mechanism and a characterisation of the optimal mechanism which minimises the maximal expected error within the class of mechanisms considered.

Acknowledgement

The first named author was supported by the Science Foundation Ireland grant SFI/11/PI/1177.

References

[1] M. Barbaro, T.J. Zeller, A face is exposed for AOL searcher no. 4417749, New York Times, Aug, 2006.
[2] J. Blocki, A. Blum, A. Datta, O. Sheffet, Differentially private data analysis of social networks via restricted sensitivity, in: Proceedings of the 4th Conference on Innovations in Theoretical Computer Science, ACM, 2013, pp. 87–96.
[3] R. Chen, N. Mohammed, B.C. Fung, B.C. Desai, L. Xiong, Publishing set-valued data via differential privacy, Proc. VLDB Endowm. 4 (11) (2011) 1087–1098.
[4] D. Dobkin, A.K. Jones, R.J. Lipton, Secure databases: Protection against user influence, ACM Trans. Database Syst. (TODS) 4 (1) (1979) 97–106.
[5] C. Dwork, Differential privacy, in: Encyclopedia of Cryptography and Security, Springer, 2011, pp. 338–340.
[6] C. Dwork, Differential privacy: A survey of results, in: Theory and Applications of Models of Computation, Springer, 2008, pp. 1–19.
[7] F.K. Dankar, K. El Emam, The application of differential privacy to health data, in: Proceedings of the 2012 Joint EDBT/ICDT Workshops, ACM, 2012, pp. 158–166.
[8] Ú. Erlingsson, V. Pihur, A. Korolova, Rappor: Randomized aggregatable privacy-preserving ordinal response, in: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, ACM, 2014, pp. 1054–1067.
[9] S.E. Fienberg, U.E. Makov, R.J. Steele, Disclosure limitation using perturbation and related methods for categorical data, J. Official Stat. 14 (4) (1998) 485–502.
[10] M. Hardt, K. Talwar, On the geometry of differential privacy, in: Proceedings of the Forty-second ACM Symposium on Theory of Computing, ACM, 2010, pp. 705–714.
[11] N. Holohan, D.J. Leith, O. Mason, Differential privacy in metric spaces: Numerical, categorical and functional data under the one roof, Inform. Sci. 305 (2015) 256–268.
[12] P. Kairouz, S. Oh, P. Viswanath, Extremal mechanisms for local differential privacy, in: Advances in Neural Information Processing Systems, 2014, pp. 2879–2887.
[13] V. Karwa, S. Raskhodnikova, A. Smith, G. Yaroslavtsev, Private analysis of graph structure, Proc. VLDB Endowm. 4 (11) (2011) 1146–1157.
[14] N. Li, T. Li, S. Venkatasubramanian, t-closeness: Privacy beyond k-anonymity and ℓ-diversity, in: Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on, IEEE, 2007, pp. 106–115.
[15] A. Machanavajjhala, D. Kifer, J. Gehrke, M. Venkitasubramaniam, ℓ-diversity: Privacy beyond k-anonymity, ACM Trans. Knowl. Discov. Data (TKDD) 1 (1) (2007) 3.
[16] F. McSherry, K. Talwar, Mechanism design via differential privacy, in: Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on, IEEE, 2007, pp. 94–103.
[17] N. Mohammed, R. Chen, B. Fung, P.S. Yu, Differentially private data release for data mining, in: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2011, pp. 493–501.
[18] A. Narayanan, V. Shmatikov, Robust de-anonymization of large sparse datasets, in: Security and Privacy, 2008. SP 2008. IEEE Symposium on, IEEE, 2008, pp. 111–125.
[19] L. Sweeney, k-anonymity: A model for protecting privacy, Int. J. Uncertain. Fuzziness Knowledge-Based Syst. 10 (05) (2002) 557–570.
[20] S.L. Warner, Randomized response: A survey technique for eliminating evasive answer bias, J. Amer. Statist. Assoc. 60 (309) (1965) 63–69.
[21] X. Zhang, X. Meng, Discovering top-k patterns with differential privacy-an accurate approach, Frontiers Of Comput. Sci. 8 (5) (2014) 816–827.