Information Sciences 305 (2015) 256–268

Contents lists available at ScienceDirect

Information Sciences

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

Differential privacy in metric spaces: Numerical, categorical and functional data under the one roof

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

[a] *School of Computer Science and Statistics, Trinity College Dublin, Ireland*
[b] *Hamilton Institute/Department of Mathematics & Statistics, Maynooth University-National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland*

ARTICLE INFO

Article history:
Received 24 February 2014
Received in revised form 19 June 2014
Accepted 27 January 2015
Available online 3 February 2015

Keywords:
Differential privacy
Metric space
Categorical data
Functional data
Data sanitisation

ABSTRACT

We study differential privacy in the abstract setting of probability on metric spaces. Numerical, categorical and functional data can be handled in a uniform manner in this setting. We demonstrate how mechanisms based on data sanitisation and those that rely on adding noise to query responses fit within this framework. We prove that once the sanitisation is differentially private, then so is the query response for any query. We show how to construct sanitisations for high-dimensional databases using simple 1-dimensional mechanisms. We also provide lower bounds on the expected error for differentially private sanitisations in the general metric space setting. Finally, we consider the question of sufficient sets for differential privacy and show that for relaxed differential privacy, any algebra generating the Borel σ-algebra is a sufficient set for relaxed differential privacy.

1. Introduction

1.1. Background

The rapid expansion of the Internet and its use in everyday life, alongside the growing understanding of the potential benefits of big data [5], has pushed data privacy to the forefront of research priorities since the turn of the millennium. Whether it be online, in the supermarket or at the hospital, corporations and governments are collecting vast quantities of data about our activities, the choices we make and the people we are in order to work more efficiently, increase profits and better serve our needs as consumers and citizens [7]. The challenge of making this potentially highly sensitive data publicly available where it can be put to good use is far from trivial and it is with this problem that the field of data privacy is concerned.

Various researchers and practitioners have considered applying anonymisation techniques to data sets such as removing explicit identifiers (name, address, telephone number, social security number, etc.) while leaving quasi-identifiers[1] in place. While these anonymised data sets do indeed preserve participants' privacy in isolation, auxiliary/background information can make this technique extremely vulnerable to attack [19]. A study by Sweeney in 2000 [25] found that as much as 87% of the US population (216 out of 248 million people) could be uniquely identified using only three quasi-identifiers (5-digit ZIP code, gender and date of birth). This meant census data could be linked to "anonymised" health records to determine the health status of unsuspecting patients.

* Corresponding author. Tel.: +353 (0)1 7086274; fax: +353 5(0)1 7086269.
E-mail address: oliver.mason@nuim.ie (O. Mason).

[1] A quasi-identifier is an attribute that is not sufficient to identify an individual by itself, but can do so when combined with other quasi-identifiers (*e.g.* gender, date of birth, etc.).

http://dx.doi.org/10.1016/j.ins.2015.01.021
0020-0255/© 2015 Elsevier Inc. All rights reserved.

Then in 2006, American media firm AOL released 20 million Internet search queries, with user numbers in place of other quasi-identifiers to protect users' identities. This shield of anonymity was not sufficient for privacy to be protected however, and the data was quickly removed from the public domain [2]. Similarly in 2008, Narayanan and Shmatikov [20] successfully de-anonymised entries in an anonymised data set containing movie ratings of 500,000 subscribers which was released by the movie streaming website Netflix. The authors used the publicly-available Internet Movie Database as background information and were able to positively identify known users despite the absence of explicit identifiers in the Netflix data set. Anonymisation methods such as k-anonymity [26] and l-diversity [19] have been shown to be vulnerable to attacks based on background information [19,18].

The work discussed above underlines the unsatisfactory nature of ad hoc privacy solutions and the need for a solid theoretical foundation for privacy research. With this in mind, the concept of differential privacy was proposed in [9] to provide a formal, mathematical framework for analysing privacy-preserving data publishing and mining. The premise of differential privacy is that the outputs of queries to a database are unlikely to change substantially with the addition of a new participant's information. This means that outputs will be similar whether or not an individual participates in the database.

There is now a considerable body of work on differential privacy in the theoretical Computer Science literature [10]. Many of the papers in the literature concern data or queries of some particular type or on the development of particular algorithms that satisfy differential privacy. For instance, the design of differentially private algorithms for calculating singular vectors is considered in [16], while differentially private recommender systems are developed in [23]; in both of these instances, the data are naturally modelled as real numbers. Algorithms for search problems and learning are considered in [3,17]. A statistical perspective on differential privacy was developed in [27]; this paper considered real-valued ($[0,1]$ in fact) queries and data. In [13], mechanisms that maximise a suitable utility function were investigated; this paper assumed discrete finite-valued data spaces, which can describe categorical data. The recent paper [24] addressed the design of optimal mechanisms that add noise independently of the data; the queries considered are real-valued. To date, the only major reference on differential privacy for functional data appears to be [14]; in the same paper the authors emphasise the importance of being careful in selecting the measure space with respect to which probabilities are defined. In particular, if we choose our σ-algebra to be the trivial one consisting of the empty set and the entire space, then every mechanism is differentially private. The work of [15] and other similar papers on lower bounds for differentially private mechanisms considers real (or in some cases integer) valued data.

The principal aim in this paper is to develop a unifying, abstract framework for all major data and query types considered so far. It is not our intention to add to the considerable body of work done on differential privacy for specific data types or on output perturbations for specific query types (counting queries, linear queries, etc.). Rather, we identify the minimal technical requirements necessary for a discussion of differential privacy and initiate a programme of developing a theory based on these. Such an approach has a number of advantages. Isolating the core assumptions behind results provides insight into precisely why they hold and in some instances, proofs are clarified. Moreover results developed at an abstract level can be widely applied as their derivation is not tied to any one particular application domain. Throughout the paper, we emphasise this point with examples of real-valued, functional and categorical data.

Metrics have previously been used in the context of differential privacy in the papers [12,6], albeit in a different manner to that employed here. The standard definition of differential privacy (see Section 2.5 for details) measures similarity between databases using *Hamming distance*: namely the number of places or entries in which they differ. The work of these earlier papers considers generalisations of this in which arbitrary metrics may be used to quantify similarity of databases. It is worth noting that the results of Sections 3 and 4 may be readily extended to more general measures of similarity in the same spirit. However, our interest lies in analysing the standard definition of relaxed differential privacy and metrics are used here to quantify the error of a mechanism.

1.2. Our results

The principal contributions of this paper are the following.

- We consider differential privacy in the general framework of probability on metric spaces and highlight with specific examples of practical relevance that it can be seamlessly applied to numerical, categorical and functional data. A major advantage of such an abstract framework is that results derived in generality can be applied uniformly to a wide variety of different applications.
- Our description shows how mechanisms based on database sanitisations and output perturbations to query responses can be treated in a unified fashion. Moreover, the fundamental differences between these two classes of mechanisms can be seen clearly within the framework developed here (see the Remark after Theorem 4).
- We describe techniques for generating families of (ϵ,δ) differentially private mechanisms from simpler mechanisms. One such example is given by sanitised response mechanisms generated from an (ϵ,δ) differentially private database sanitisation. We also show how to generate differentially private sanitisations for high-dimensional databases using sanitisations for 1-dimensional databases. This approach provides a simple paradigm that can be applied to any type of data.

- We describe lower bounds for the error in releasing a database in an (ϵ, δ)-differentially private fashion using product sanitisations. This result applies to data drawn from any metric space in contrast to previous work, which has largely focussed on real and integer valued data.
- We consider the question of testing differential privacy and describe sufficient sets for (ϵ, δ) differential privacy. This work could be useful in the design of tests for differential privacy as it significantly reduces the size of the class of sets that must be considered.

1.3. Structure of paper

We begin in Section 2 by establishing the measure-theoretic framework for differential privacy. We then consider the question of sufficient sets for differential privacy in Section 3, address sanitised response mechanisms in Section 4 and focus on product sanitisations in Section 5. Section 6 considers accuracy and we give concluding remarks in Section 7.

2. Preliminaries

We first recall some standard concepts and results from probability and measure theory [4,22]. Given an algebra $\mathcal{S}$ of subsets of a set Ω, we use $\sigma(\mathcal{S})$ to denote the smallest σ-algebra containing $\mathcal{S}$ and refer to $\sigma(\mathcal{S})$ as the σ-algebra generated by $\mathcal{S}$. A set Ω together with a σ-algebra of subsets of Ω is a measurable space.

Given a mapping $Q : U \to E$ from a set U to a set E and a subset $A \subseteq E$, the notation $Q^{-1}(A)$ denotes the pre-image of $A, Q^{-1}(A) = \{u \in U : Q(u) \in A\}$.

A monotone class $\mathcal{M}$ of subsets of some set Ω is defined by the following two properties: (i) if $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$, and if $A_i \subseteq A_{i+1}$ for all i, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$; (ii) if $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$, and if $A_i \supseteq A_{i+1}$ for all i, then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}$.

The next result, which appears as Theorem 3.4 in [4], characterises $\sigma(\mathcal{S})$ as the smallest monotone class containing $\mathcal{S}$.

Theorem 1. *Let $\mathcal{S}$ be an algebra of subsets of some set Ω and let $\mathcal{M}$ be a monotone class such that $\mathcal{S} \subseteq \mathcal{M}$. Then $\sigma(\mathcal{S}) \subseteq \mathcal{M}$.*

Given two measurable spaces $(X, \mathcal{A}_X)$ and $(Y, \mathcal{A}_Y)$, subsets of $X \times Y$ of the form

$$R = \bigcup_{i=1}^{p} X_i \times Y_i,$$

where $X_i \in \mathcal{A}_X, Y_i \in \mathcal{A}_Y$ for $1 \leqslant i \leqslant p$ and $(X_i \times Y_i) \cap (X_j \times Y_j) = \emptyset$ for $i \neq j$ are known as *elementary subsets*. Let $\mathcal{R}$ denote the collection of all elementary subsets and denote the usual product σ-algebra on $X \times Y$ by $\mathcal{A}_{X \times Y}$. The following result is Theorem 8.3 of [22].

Theorem 2. *If $\mathcal{M}$ is a monotone class and $\mathcal{R} \subseteq \mathcal{M}$, then $\mathcal{A}_{X \times Y} \subseteq \mathcal{M}$.*

Finally, for a measure μ on a measurable space $(X, \mathcal{A}_X)$, we recall the following simple fact.

Proposition 1. *Suppose that $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}_X$ satisfies $A_i \subseteq A_{i+1}$ for all i, then $\lim_{i \to \infty} \mu(A_i) = \mu(\bigcup_{i=1}^{\infty} A_i)$.*

Similarly, if $A_i \supseteq A_{i+1}$ for all i, then $\lim_{i \to \infty} \mu(A_i) = \mu(\bigcap_{i=1}^{\infty} A_i)$.

2.1. Database model

The individual entries of the databases we consider are elements of a set $D \subseteq U$ where U is a metric space with metric ρ. We equip U with the Borel σ-algebra generated by the open sets in U (in the metric topology); D then naturally inherits a σ-algebra $\mathcal{A}_D$. A database $\mathbf{d}$ with n rows is given by a vector $\mathbf{d} = (d_1, \ldots, d_n) \in D^n$ in which $d_i \in D$ is the ith row. Throughout, we assume that U^n (and D^n) is equipped with the usual product σ-algebra $\mathcal{A}_{U^n}$ generated by $\{A_1 \times \cdots \times A_n : A_i \in \mathcal{A}_U\}$. This ensures that projection maps $\pi_i : U^n \to U$ given by $\pi_i(x_1, \ldots, x_n) = x_i$ are measurable.

It is worth highlighting the generality of this setting: the metric space D can contain numerical, categorical or functional data; moreover, it can be discrete or continuous.

Example 1. If our data concern the hobbies or interests of people, we consider a set of all possible hobbies, denoted by $\mathcal{H}$. For simplicity it is not unreasonable to assume that $\mathcal{H}$ is finite. Our data entries are then drawn from the power set $D := 2^{\mathcal{H}}$ of $\mathcal{H}$, which will again be a finite set. There are various natural choices of metric in this case. We could consider the discrete metric on D in which $\rho_1(A, B) = 1$ if $A \neq B$ and 0 otherwise. Alternatively, we could choose the metric given by symmetric distance: $\rho_2(A, B) = |(A \cup B) \backslash (A \cap B)|$. In both of these cases, the Borel σ-algebra consists of all subset of D. Note that there is no requirement that each entry in a database in D^n have the same size or cardinality, reflecting the fact that not all of us have the same number of interests or hobbies.

Example 2. In readings from field deployed sensors, each reading has a time-stamp giving rise to time-course data. Another example is in smart metering where the supplier collects data from consumers giving electricity consumption over a time-window. Data of this type is naturally represented as either a function or a sequence of real numbers. In our framework, we can take U to be a sequence space such as l_∞ or l_2, or an appropriate function space such as $C([0,T])$ or $L_2([0,T])$, where T represents the billing period (for instance). All of these spaces have natural norms defined on them (in fact they are all Banach spaces) and can be equipped with the Borel σ-algebra generated from the open sets in the norm topology.

We say that two databases $\mathbf{d} = (d_1, \ldots, d_n)$ and $\mathbf{d}' = (d'_1, \ldots, d'_n)$ in D^n are *neighbours*, and write $\mathbf{d} \sim \mathbf{d}'$, if there is some $j \in \{1, \ldots, n\}$ such that $d_j \neq d'_j$ and $d_i = d'_i$ for all $i \in \{1, \ldots, n\} \setminus \{j\}$. More generally, we denote by $h(\mathbf{d}, \mathbf{d}')$ the *Hamming distance* between $\mathbf{d}$ and $\mathbf{d}'$.

For the most part, we assume that the data space D is compact. This is immediate if D is finite (as in Example 1) and is a natural assumption in most realistic situations. When D is compact, we denote by $\mathrm{diam}(D)$ the diameter of D:

$$\mathrm{diam}(D) = \max_{d,d' \in D} \rho(d, d'). \tag{1}$$

2.2. Query model

We consider a very general query model. The set of all possible responses is denoted by E_Q while $\mathcal{A}_Q$ is a σ-algebra of subsets of E_Q. In the case where E_Q is a metric space, we assume that $\mathcal{A}_Q$ is the Borel σ-algebra generated by the metric topology. A query $Q(\mathbf{d})$ is then a measurable function, $Q : U^n \to E_Q$ and hence $Q^{-1}(A) \in \mathcal{A}_{U^n}$ for all $A \in \mathcal{A}_Q$.

Example 3. As with the data in $\mathbf{d}$, queries are not restricted to take numerical values in this setting. For instance, if we consider Example 1 above, then we could consider a query asking for the number of people in the database who are interested in Classical Music or Football for instance: this would clearly be a numeric query. On the other hand, we could also request the 3 most common hobbies in the database, the output of which would be a set.

We next formally introduce the concept of a *response mechanism* within this general framework. As is standard in the literature on probability theory, we assume that a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given [4]. Here, Ω is a set (corresponding informally to the set of outcomes of some 'random experiment'), $\mathcal{F}$ is a σ-algebra of subsets of Ω (corresponding to 'events' of interest) while $\mathbb{P}$ is a probability measure defined on $\mathcal{F}$.

We define a response mechanism with respect to a set of queries, $\mathcal{Q}$ which represent the queries of interest for a particular application. A response mechanism with respect to this set is a collection of measurable mappings

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

Note that $X_{Q,\mathbf{d}}$ is an E_Q-valued random variable for each Q and $\mathbf{d}$. This naturally reflects the fact that a response mechanism is typically described as a random algorithm that is indexed by the query Q and the database $\mathbf{d}$. We remind the reader that $\mathbb{P}(X_{Q,\mathbf{d}} \in A)$ is standard shorthand in probability theory for $\mathbb{P}(\{\omega \in \Omega : X_{Q,\mathbf{d}}(\omega) \in A\})$. Throughout the paper, in cases where there is no ambiguity, the argument ω of $X_{Q,\mathbf{d}}(\omega)$ will be suppressed.

2.3. Sanitised response mechanisms

Much of our work in this paper is focused on a particular type of mechanism, the sanitised response mechanism. First of all, we introduce the notation of a *sanitisation* which is a family of measurable mappings

$$\{Y_\mathbf{d} : \Omega \to U^n | \mathbf{d} \in D^n\}.$$

This represents "sanitising" the original database. If the database is sanitised by adding appropriate noise, the mapping $Y_\mathbf{d}(\omega)$ takes the form $Y_\mathbf{d}(\omega) = \mathbf{d} + N(\omega)$ for some U^n-valued random vector $N(\omega)$. Note that in order to define a mechanism by adding noise, it is necessary for U^n to have a suitable algebraic structure. However, our framework does not require this extra structure to be present.

A response mechanism $X_{Q,\mathbf{d}}$ is said to be a *sanitised response mechanism* if $X_{Q,\mathbf{d}}$ takes the particular form

$$X_{Q,\mathbf{d}} = Q \circ Y_\mathbf{d} \tag{3}$$

for all $Q \in \mathcal{Q}$ and $\mathbf{d} \in D^n$.

The motivation behind this choice of terminology is that the mechanism is generated by first sanitising the database via the sanitisation $Y_\mathbf{d}$ and then answering a query Q using the sanitised database. That is, by answering with $Q(Y_\mathbf{d}(\omega))$ rather than the true query answer $Q(\mathbf{d})$.

2.4. Output perturbations

In the differential privacy literature, output perturbation mechanisms which operate by perturbing or adding noise to the query response have attracted particular interest [9]. Such mechanisms can also be expressed in the general framework described here. Let a query $Q : U^n \to E_Q$ be given. Assume that there is a family of measurable functions $\{Z_q : \Omega \to E_Q | q \in E_Q\}$ defined on the output space E_Q of Q. Informally, for $q \in E_Q, Z_q$ can be thought of as the random response of the mechanism when the true query response is q. An output perturbation mechanism is then defined as a response mechanism $X_{Q,\mathbf{d}}$ of the particular form

$$X_{Q,\mathbf{d}} = Z_{Q(\mathbf{d})}. \tag{4}$$

For real-valued data, the functions Z_q may take the form $Z_q(\omega) = q + N(\omega)$ where N represents the noise added to the query response as in the well known Laplacian mechanism. For discrete-valued queries, Z_q can be defined by specifying an appropriate probability mass function on E_Q.

2.5. Differential privacy

In the interest of completeness and clarity, we now recall the definition of differential privacy and write it in the setting of this paper.

Definition 1 (*Differential privacy with respect to a query*). Let $\epsilon \geqslant 0, 0 \leqslant \delta \leqslant 1$ be given. A response mechanism is (ϵ,δ)-differentially private with respect to a query $Q_0 \in \mathcal{Q}$ if for all $\mathbf{d} \sim \mathbf{d}' \in D^n$ and all $A \in \mathcal{A}_{Q_0}$,

$$\mathbb{P}(X_{Q_0,\mathbf{d}} \in A) \leqslant e^{\epsilon}\mathbb{P}(X_{Q_0,\mathbf{d}'} \in A) + \delta. \tag{5}$$

It is important to note that the relation $\mathbf{d} \sim \mathbf{d}'$ is symmetric, so inequality (5) is required to hold when $\mathbf{d}$ and $\mathbf{d}'$ are swapped.

Definition 2 (*Differential privacy*). A response mechanism is (ϵ,δ)-differentially private with respect to a set of queries $\mathcal{Q}$ if it is (ϵ,δ)-differentially private with respect to every query $Q_0 \in \mathcal{Q}$.

The above definitions are often referred to as *relaxed* differential privacy; the original notion of differential privacy introduced in [9] considers the case where $\delta = 0$.

Throughout the paper, when discussing sanitised response mechanisms, we shall say that the sanitisation $Y_{\mathbf{d}}$ is (ϵ,δ)-differentially private if $\mathbb{P}(Y_{\mathbf{d}} \in A) \leqslant e^{\epsilon}\mathbb{P}(Y_{\mathbf{d}'} \in A) + \delta$, for all $\mathbf{d} \sim \mathbf{d}'$ and $A \in \mathcal{A}_{U^n}$.

3. Sufficient sets for differential privacy

In this brief section, we consider the following question: is there a strict subset $\mathcal{S}$ of $\mathcal{A}_Q$ such that if (5) is satisfied for all A in $\mathcal{S}$, it is guaranteed to be satisfied for all A in $\mathcal{A}_Q$? We refer to such a set as a sufficient set for differential privacy.

Depending on the application, the query output space may be a subset of $\mathbb{R}^n$ or of a sequence or function space such as $C([0,T])$. A key question for the practical deployment of differentially private mechanisms is how to determine if a mechanism is in fact (ϵ,δ)-differentially private. Testing (5) on the entire σ-algebra is clearly a prohibitively difficult task. Our next result shows that it is sufficient to test this condition on any *algebra* $\mathcal{S} \subset \mathcal{A}_Q$ that generates $\mathcal{A}_Q$.

Theorem 3. *Let a response mechanism (2) and a query $(Q, E_Q, \mathcal{A}_Q)$ be given and let $\mathcal{S} \subset \mathcal{A}_Q$ be an algebra such that $\sigma(S) = \mathcal{A}_Q$. If (5) holds for all sets $A \in \mathcal{S}$, then it holds for all sets $A \in \mathcal{A}_Q$.*

Proof. Let $\mathcal{B}$ denote the collection of sets in E_Q for which (5) holds. By assumption, $\mathcal{S} \subseteq \mathcal{B}$. Now let $A_1, A_2, \ldots$ be any collection of sets in $\mathcal{B}$ with $A_i \subseteq A_{i+1}$ for all i. Define $\overline{A} := \cup_i A_i$ and let $\mathbf{d}, \mathbf{d}' \in D^n$ with $\mathbf{d} \sim \mathbf{d}'$ be given. As each $A_i \in \mathcal{B}$, it follows that

$$\mathbb{P}(X_{Q,\mathbf{d}} \in A_i) \leqslant e^{\epsilon}\mathbb{P}(X_{Q,\mathbf{d}'} \in A_i) + \delta$$

for all i. As the sequence A_i is increasing, it now follows from Proposition 1 that

$$\mathbb{P}(X_{Q,\mathbf{d}} \in \overline{A}) = \lim_{i\to\infty}\mathbb{P}(X_{Q,\mathbf{d}} \in A_i) \leqslant e^{\epsilon}\lim_{i\to\infty}\mathbb{P}(X_{Q,\mathbf{d}'} \in A_i) + \delta = e^{\epsilon}\mathbb{P}(X_{Q,\mathbf{d}'} \in \overline{A}) + \delta,$$

and so $\overline{A} \in \mathcal{B}$. An identical argument shows that for any sequence $\{A_i\}$ of sets in $\mathcal{B}$ with $A_i \supseteq A_{i+1}$ for all i, and $\overline{A} = \cap_i A_i, \overline{A} \in \mathcal{B}$. Taken together these two observations imply that $\mathcal{B}$ is a *monotone class*. Moreover, $\mathcal{S} \subseteq \mathcal{B}$. The result now follows immediately from Theorem 1. □

In the next example we show how Theorem 3 can be applied to differentially private mechanisms for functional data to obtain results such as those described in Section 3.1 of [14].

Example 4. Suppose our query Q takes values in the space $C([0,1])$ of continuous functions on $[0,1]$ equipped with the norm $\|f\|_\infty = \sup\{|f(t)| : t \in [0,1]\}$ and the σ-algebra $\mathcal{A}_Q$ of Borel sets generated by the norm topology. Let a mechanism $X_{Q,\mathbf{d}}$ be given. Then $X_{Q,\mathbf{d}}(\omega)$ is in $C([0,1])$ for each $\omega \in \Omega$.

Given a positive integer k and real numbers $0 \leqslant t_1 < \cdots < t_k \leqslant 1$, consider the projection $\pi_{t_1,\ldots,t_k} : C([0,1]) \to \mathbb{R}^k$ given by

$$\pi_{t_1,\ldots,t_k}(f) = (f(t_1),\ldots,f(t_k)).$$

These mappings are measurable with respect to the usual Borel σ-algebra on $\mathbb{R}^k$ and hence we can define the $\mathbb{R}^k$-valued mechanism $X^{t_1,\ldots,t_k}_{Q,\mathbf{d}} = \pi_{t_1,\ldots,t_k} \circ X_{Q,\mathbf{d}}$.

We claim that if the finite-dimensional mechanisms $X^{t_1,\ldots,t_k}_{Q,\mathbf{d}}$ are (ϵ,δ)-differentially private for all k and $t_1,\ldots,t_k$, then the mechanism $X_{Q,\mathbf{d}}$ is (ϵ,δ)-differentially private. The argument to show this is as follows. From the assumption on the finite-dimensional mechanisms, it follows immediately that (5) holds for all (so-called cylinder sets) sets A of the form

$$A = \pi^{-1}_{t_1,\ldots,t_k}(B)$$

where B is a Borel set in $\mathbb{R}^k$. These sets form an algebra and it follows from Theorem VII.2.1 (page 212) of [21] that the σ-algebra they generate is the Borel σ-algebra of $C([0,1])$. It follows immediately from Theorem 3 that $X_{Q,\mathbf{d}}$ defines an (ϵ,δ)-differentially private mechanism on $C([0,1])$ as claimed.

4. Sanitised response mechanisms and the identity query

A popular approach to designing differentially private response mechanisms is to add *noise* to the query response $Q(\mathbf{d})$. It is known however, that this can lead to privacy compromises by averaging a large number of responses to an identical query [10], unless the number or type of queries that can be asked is restricted. We now show that if a sanitised response mechanism (3) is (ϵ,δ)-differentially private with respect to the identity query, then it is (ϵ,δ)-differentially private with respect to *any* query. It is important to appreciate that we place minimal restrictions on the query Q and its output space. For example, if we consider E_Q to be the space of square summable or bounded real sequences then Q naturally corresponds to an infinite sequence of individual real-valued queries, each of which is a projection of the given Q.

The identity query I_n is defined by the identity map on the ambient space U^n. Formally, $I_n(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in U^n$. Note that an output space E_Q is associated with each query and that these spaces can be different for different queries; in the case of the identity query we select $E_Q = U^n$. In sanitised response mechanisms, the "sanitised database" $Y_{\mathbf{d}}(\omega)$ can be viewed as a response to the identity query. However, disclosing $Y_{\mathbf{d}}(\omega)$ need *not* disclose the original database $\mathbf{d}$ provided that an appropriate privacy-preserving perturbation has been applied. Importantly, if (ϵ,δ)-differential privacy is achieved with respect to I_n, then the response to *any* query is (ϵ,δ)-differentially private.

Theorem 4 (Identity query). *Consider a sanitised response mechanism as defined in (3). Suppose that the sanitisation $Y_{\mathbf{d}}$ is (ϵ,δ)-differentially private. Then the mechanism (3) is (ϵ,δ)-differentially private with respect to any query $(Q, E_Q, \mathcal{A}_Q)$.*

Proof. Let $\mathbf{d}, \mathbf{d}'$ be two elements of D^n with $\mathbf{d} \sim \mathbf{d}'$. By assumption,

$$\mathbb{P}(Y_{\mathbf{d}} \in E) \leqslant e^\epsilon \mathbb{P}(Y_{\mathbf{d}'} \in E) + \delta, \tag{6}$$

for all $E \in \mathcal{A}_{U^n}$. Let a query $(Q, E_Q, \mathcal{A}_Q)$ and $A \in \mathcal{A}_Q$ be given. As Q is measurable, $Q^{-1}(A) \in \mathcal{A}_{U^n}$. Then, using (6),

$$\mathbb{P}(X_{Q,\mathbf{d}} \in A) = \mathbb{P}(Q(Y_{\mathbf{d}}) \in A) = \mathbb{P}(Y_{\mathbf{d}} \in Q^{-1}(A)) \leqslant e^\epsilon \mathbb{P}\left(Y_{\mathbf{d}'} \in Q^{-1}(A)\right) + \delta = e^\epsilon \mathbb{P}(Q(Y_{\mathbf{d}'}) \in A) + \delta.$$

Thus, the mechanism satisfies (ϵ,δ)-differential privacy with respect to Q also. □

Remark. The previous result shows that the following natural intuition is valid in the abstract setting described here. If we can release a sanitised version of the database in an (ϵ,δ)-differentially private manner, then we can answer any query on this sanitisation in an (ϵ,δ)-differentially private manner also. To highlight a fundamental difference between sanitised response mechanisms and those based on output perturbations, consider the scenario where the same query is asked multiple (k) times. We can easily model this as a single query $Q^{(k)}$ which maps from U^n to $E^k_Q = E_Q \times E_Q \times \ldots \times E_Q$ (the k-fold direct product of E_Q) and is given by $Q^{(k)}(\mathbf{d}) = (Q(\mathbf{d}), Q(\mathbf{d}), \ldots, Q(\mathbf{d}))$. Theorem 4 shows that if the sanitisation $Y_{\mathbf{d}}$ is (ϵ,δ)-differentially private then the sanitised response mechanism given by $Q^{(k)} \circ Y_{\mathbf{d}}$ is also (ϵ,δ)-differentially private for all $k \geqslant 1$.

In contrast, consider the situation for output perturbation mechanisms. We will see that the conditions for (ϵ,δ)-differential privacy can be violated easily for repeated queries when using output perturbation mechanisms that are (ϵ,δ)-differentially private for a single query. It is well known that there are issues with output perturbation mechanisms in this regard and that, in particular, by averaging the responses to different queries inferences can be made concerning private data. Our purpose here is simply to show that within our formalism, answering the same query multiple times with an output perturbation mechanism can violate the defining inequality (5).

For simplicity suppose we have a binary-valued query $Q : U^n \to \{0,1\}$. So $E_Q = \{0,1\}$ and to define the output perturbation, we need to specify the distributions of Z_0 and Z_1. If we put $\mathbb{P}(Z_i = i) = 1-p, \mathbb{P}(Z_i \neq i) = p$ for $i = 0,1$, then it is not difficult to verify that the output perturbation mechanism $X_{Q,\mathbf{d}} = Z_{Q(\mathbf{d})}$ is (ϵ,δ)-differentially private if and only if

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

Let us make the reasonable assumption that there exist two neighbouring databases $\mathbf{d},\mathbf{d}'$ in D^n for which the response to Q is different; say $Q(\mathbf{d}) = 0, Q(\mathbf{d}') = 1$. Then for the set $A = \{0\}$, we have $\mathbb{P}(Z_{Q(\mathbf{d})} \in A) = \mathbb{P}(Z_0 = 0) = 1-p$, and similarly $\mathbb{P}(Z_{Q(\mathbf{d}')} \in A) = p$.

Now suppose the query Q is asked twice and we wish to use our output perturbation mechanism to answer it privately. The output space is $E_Q \times E_Q$ and the natural way to define the random variables $Z_{(q_1,q_2)}$ for $q_1,q_2 \in E_Q$ is by setting $Z_{(q_1,q_2)} = (Z_1,Z_2)$ where the Z_i are independent and Z_i is identically distributed to Z_{q_i} for $i = 1,2$. For the scenario described in the last paragraph, if we choose $\epsilon = 0.1$ and $\delta = 0.4$, then if we choose $p = 0.286$ the mechanism is (ϵ,δ)-differentially private with respect to Q. However if we repeat the query twice and consider the set $A = \{0\}$ as before,

$$\mathbb{P}\left(Z^{(2)}_{Q(\mathbf{d})} \in A \times A\right) = (1-p)^2 = 0.5098,$$

while

$$\mathbb{P}\left(Z^{(2)}_{Q(\mathbf{d}')} \in A \times A\right) = p^2 = 0.081796.$$

It is now straightforward to verify by direct calculation that

$$\mathbb{P}\left(Z^{(2)}_{Q(\mathbf{d}_1)} \in A \times A\right) > e^{\epsilon}\mathbb{P}\left(Z^{(2)}_{Q(\mathbf{d}_2)} \in A \times A\right) + \delta$$

for $\epsilon = 0.1, \delta = 0.4$ showing that applying the output perturbation mechanism twice in this instance will lead to differential privacy being broken.

Corollary 1. *Consider a sanitised response mechanism as defined in (3) and suppose $I_n \in \mathcal{Q}$. Then this response mechanism is (ϵ,δ)-differentially private with respect to I_n if and only if it is (ϵ,δ)-differentially private with respect to every query $Q \in \mathcal{Q}$.*

Proof. "⇒": Theorem 4."⇐": The response mechanism is (ϵ,δ)-differentially private with respect to every query $Q \in \mathcal{Q}$ by assumption, therefore it must be (ϵ,δ)-differentially private with respect to the identity query I_n, since $I_n \in \mathcal{Q}$. □

5. Product sanitisations

In this section, we derive a result that relates differentially private sanitised response mechanisms for n-dimensional databases in D^n to mechanisms for simple 1-dimensional databases.

Before showing how differentially private sanitised response mechanisms for databases in D^n can be constructed from simple mechanisms for databases in D, we first establish a number of technical results.

Lemma 1. *Let $A_1,\ldots,A_p,B_1,\ldots,B_p$ be two collections of non-empty sets. Then the finite union $\bigcup_{i=1}^p(A_i \times B_i)$ can be written as*

$$\bigcup_{i=1}^{p}(A_i \times B_i) = \bigcup_{I \subseteq \{1,\ldots,p\}} (\widetilde{A}_I \times \widetilde{B}_I),$$

where $\widetilde{A}_I = \bigcup_{i\in I}A_i$ and $\widetilde{B}_I = \bigcap_{i\in I}B_i \backslash \bigcup_{i\notin I}B_i$. Moreover, $\widetilde{B}_I \cap \widetilde{B}_J = \emptyset$ for all $I \neq J$.

Proof. We need to prove equality and disjointness of the decomposition. Let $(a,b) \in \bigcup_{i=1}^p(A_i \times B_i)$. Then there exists at least one i^* such that $(a,b) \in A_{i^*} \times B_{i^*}$. Let $I_b := \{i : b \in B_i\} \subseteq \{1,\ldots,p\}$ (note $i^* \in I_b$). Then $b \in \bigcap_{i\in I_b}B_i$, but $b \notin B_j$ for any $j \notin I_b$, otherwise j would be an element of I_b. Hence $b \in \bigcap_{i\in I_b}B_i \backslash \bigcup_{i\notin I_b}B_i$. Also $a \in \bigcup_{i\in I_b}A_i$ since $a \in A_{i^*}$. Hence $(a,b) \in \bigcup_{I\subseteq\{1,\ldots,p\}}\left(\bigcup_{i\in I}A_i \times \bigcap_{i\in I}B_i \backslash \bigcup_{i\notin I}B_i\right)$.

Let $(a,b) \in \bigcup_{I \subseteq \{1,\ldots,p\}} \left(\bigcup_{i\in I} A_i \times \bigcap_{i\in I} B_i \setminus \bigcup_{i\notin I} B_i\right)$. Then there exists at least one $I^* \subseteq \{1,\ldots,p\}$ such that $(a,b) \in \bigcup_{i\in I^*} A_i \times \bigcap_{i\in I^*} B_i \setminus \bigcup_{i\notin I^*} B_i$. Hence $a \in A_i$ for at least one $i \in I^*$ and $b \in B_i$ for all $i \in I^*$ and so there exists at least one $i \in I^*$ such that $(a,b) \in A_i \times B_i$ and so $(a,b) \in \bigcup_{i=1}^{p}(A_i \times B_i)$.

Finally, we show that $\widetilde{B}_I \cap \widetilde{B}_J = \emptyset$ if $I \neq J$. To see this, note that if $I \neq J$, then we can without loss of generality assume that there is some index $k \in I$ that is not in J. Then any $x \in \widetilde{B}_I$ must be in B_k. However, as $k \in J^C$, it follows that $x \in \cup_{i\notin J} B_k$ and hence that $x \notin B_J$. This shows that the intersection is empty as claimed. □

For future use, we note that an analogous argument to that given above can be used to show the following.

Lemma 2. *Let $A_1,\ldots,A_p,B_1,\ldots,B_p$ be two collections of non-empty sets. Then the finite union $\bigcup_{i=1}^{p}(A_i \times B_i)$ can be written as*

$$\bigcup_{j=1}^{q}\left(\widetilde{A}_j \times \widetilde{B}_j\right),$$

where $\widetilde{A}_i \cap \widetilde{A}_j = \emptyset$ for $i \neq j$.

For the remainder of this section we consider a special form for the database sanitisation $Y_{\mathbf{d}}(\omega)$. Suppose a family $\{Y_d : \Omega \to U | d \in D\}$ of measurable mappings is given. Define the mechanism $Y_{\mathbf{d}} : \Omega \to U^n$ for every $\mathbf{d} = (d_1,\ldots,d_n) \in D^n$ by

$$Y_{\mathbf{d}}(\omega) = \left(Y_{\mathbf{d}}^1(\omega),\ldots,Y_{\mathbf{d}}^n(\omega)\right), \tag{7}$$

for $\omega \in \Omega$. Here the $Y_{\mathbf{d}}^i$ are independent and $Y_{\mathbf{d}}^i$ has the same distribution as Y_{d_i}, for all $\mathbf{d} \in D^n, i \in \{1,\ldots,n\}$.

Remark. The overall sanitisation $Y_{\mathbf{d}}$ is constructed componentwise: each component is itself a 'random response' taking values in U and the individual components are independent of each other. Finally, the distribution of the random response for the i-th component is determined by the value of d_i, the i-th component of $\mathbf{d}$. For instance, if the database is real-valued then one way of implementing such a sanitisation would be to add independent and identically distributed noise to each entry of the database.

We first note the following lemma concerning such mechanisms.

Lemma 3. *Let a family $\{Y_d : \Omega \to U | d \in D\}$ of measurable mappings be given and let $Y_{\mathbf{d}}$ be defined by (7). If $Y_{\mathbf{d}}$ is (ϵ,δ)-differentially private then*

$$\mathbb{P}(Y_d \in A) \leqslant e^{\epsilon}\mathbb{P}(Y_{d'} \in A) + \delta,$$

for all $d, d' \in D, A \in \mathcal{A}_U$.

Proof. Let d, d' in D be given. If $d = d'$, the result is trivial. If $d \neq d'$, take $\mathbf{d} = (d, d_2,\ldots,d_n), \mathbf{d}' = (d', d_2,\ldots,d_n)$ for any choice of $d_2,\ldots,d_n$ in D. As $Y_{\mathbf{d}}$ is (ϵ,δ)-differentially private and the projection $\pi_1 : U^n \to U$ onto the first coordinate is measurable, it follows that for $A \in \mathcal{A}_U$:

$$\mathbb{P}(Y_d \in A) = \mathbb{P}(\pi_1(Y_{\mathbf{d}}) \in A) = \mathbb{P}(Y_{\mathbf{d}} \in \pi_1^{-1}(A)) \leqslant e^{\epsilon}\mathbb{P}(Y_{\mathbf{d}'} \in \pi_1^{-1}(A)) + \delta = e^{\epsilon}\mathbb{P}(Y_{d'} \in A) + \delta. \quad \square$$

We next note that the converse of this result also holds.

Theorem 5. *Consider a family $\{Y_d : \Omega \to U | d \in D\}$ of measurable mappings and assume that*

$$\mathbb{P}(Y_d \in A) \leqslant e^{\epsilon}\mathbb{P}(Y_{d'} \in A) + \delta,$$

for all $d, d' \in D, A \in \mathcal{A}_U$. Let $Y_{\mathbf{d}}$ be as defined in (7). Then,

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

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

Proof. We shall use induction on n. By assumption, the result is true for $n = 1$. Let $n > 1$ be given and assume that the result is true for all $k \leqslant n - 1$.

Assume that $\mathbf{d}$ and $\mathbf{d}'$ differ in the first element so $d_1 \neq d_1'$ but $d_j = d_j'$ for $j \neq 1$. Let

$$R = \bigcup_{i=1}^{p}(A_i \times B_i), \tag{8}$$

where $A_i \in \mathcal{A}_U, B_i \in \mathcal{A}_{U^{n-1}}$, be given. It follows from Lemma 1 that we can write

$$R = \bigcup_{i=1}^{q} (\widetilde{A}_i \times \widetilde{B}_i), \tag{9}$$

where $\widetilde{A}_i \in \mathcal{A}_U, \widetilde{B}_i \in \mathcal{A}_{U_{n-1}}$ for $1 \leqslant i \leqslant q$ and $\widetilde{B}_i \cap \widetilde{B}_j = \emptyset$ for $i \neq j$. Then, using the fact that the sets $\widetilde{B}_i$ are disjoint and the independence of the components of $Y_{\mathbf{d}}, Y_{\mathbf{d}'}$,

$$\begin{aligned}\mathbb{P}(Y_{\mathbf{d}} \in R) &= \sum_{i=1}^{q} \mathbb{P}(Y_{\mathbf{d}} \in \widetilde{A}_i \times \widetilde{B}_i) = \sum_{i=1}^{q} \mathbb{P}(Y_{d_1} \in \widetilde{A}_i)\, \mathbb{P}(Y_{(d_2,\ldots,d_n)} \in \widetilde{B}_i) \leqslant \sum_{i=1}^{q} \left(e^{\epsilon} \mathbb{P}(Y_{d_1'} \in \widetilde{A}_i) + \delta\right) \mathbb{P}(Y_{(d_2',\ldots,d_n')} \in \widetilde{B}_i) \\ &= e^{\epsilon} \sum_{i=1}^{q} \mathbb{P}\left(Y_{\mathbf{d}'} \in \widetilde{A}_i \times \widetilde{B}_i\right) + \delta \mathbb{P}\left(Y_{(d_2',\ldots,d_n')} \in \bigcup_{i=1}^{q} \widetilde{B}_i\right) \leqslant e^{\epsilon} \mathbb{P}(Y_{\mathbf{d}'} \in R) + \delta.\end{aligned}$$

If $d_1 = d_1'$, then $(d_2, \ldots, d_n) \sim (d_2', \ldots, d_n')$ and we can use Lemma 2 and the induction hypothesis to conclude that

$$\mathbb{P}(Y_{\mathbf{d}} \in R) \leqslant \sum_{i=1}^{q} \mathbb{P}\left(Y_{d_1'} \in \widetilde{A}_i\right)\left(e^{\epsilon} \mathbb{P}\left(Y_{(d_2',\ldots,d_n')} \in \widetilde{B}_i\right) + \delta\right) = e^{\epsilon} \sum_{i=1}^{q} \mathbb{P}\left(Y_{\mathbf{d}'} \in \widetilde{A}_i \times \widetilde{B}_i\right) + \delta \mathbb{P}\left(Y_{d_1'} \in \bigcup_{i=1}^{q} \widetilde{A}_i\right) \leqslant e^{\epsilon} \mathbb{P}(Y_{\mathbf{d}'} \in R) + \delta.$$

Thus for any set R of the form (8), we can conclude that

$$\mathbb{P}(Y_{\mathbf{d}} \in R) \leqslant e^{\epsilon} \mathbb{P}(Y_{\mathbf{d}'} \in R) + \delta. \tag{10}$$

In particular, (10) holds for all elementary sets $R \in \mathcal{A}_{U^n}$. A similar argument to that used in the proof of Theorem 3 shows that the collection of all sets satisfying (10) is a monotone class; furthermore this collection of subsets contains the elementary sets. The result now follows from Theorem 2. □

Remark. In the following subsection, we describe some simple applications of Theorem 5. It is worth noting that it applies to any database space $D \subseteq U$, and to *discrete* spaces in particular. For mechanisms of the form (7), it simplifies the task of testing the mechanism for differential privacy considerably. For instance, if D is a finite set with $|D|$ elements, then it is only necessary to check (5) for all $\binom{|D|}{2}$ pairs of elements of D and all $2^{|D|}$ subsets of D to ensure differential privacy on D^n. In general, we would have $n\binom{|D|}{2}|D|^{n-1}$ pairs of neighbouring elements and $2^{|D|^n}$ subsets to worry about!

5.1. Examples

Example 5. The addition of appropriately scaled Laplacian noise is now a standard approach to the design of differentially private responses. We note here how Theorem 5 combined with a simple adaptation of the arguments first given in [9] (for the case where $\delta = 0$) can be used to construct (ϵ, δ)-differentially private mechanisms for n-dimensional sanitisations with $\delta \neq 0$.

Recall that a Laplacian random variable $X : \Omega \to \mathbb{R}$ with mean zero and variance $2b^2$ has a probability density function (PDF) is given by

$$f(x) = \frac{1}{2b} e^{-\frac{|x|}{b}}.$$

Let $D \subset \mathbb{R}$ be bounded; for each $d \in D$, let $Y_d(\omega) = d + L(\omega)$ where $L : \Omega \to \mathbb{R}$ is a Laplacian random variable with mean zero and variance $2b^2$ such that

$$b \geqslant \frac{\operatorname{diam}(D)}{\epsilon - \log(1 - \delta)}.$$

The resulting sanitised response mechanism corresponding to (7) is (ϵ, δ)-differentially private for any database in D^n.

To see this, note that by Theorem 5, it is sufficient to show that

$$\int_A \frac{e^{-\frac{|x-d|}{b}}}{2b} \mathrm{dx} \leqslant e^{\epsilon} \int_A \frac{e^{-\frac{|x-d'|}{b}}}{2b} \mathrm{dx} + \delta,$$

for all $d, d' \in D, A \in \mathcal{B}(\mathbb{R})$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on the real line $\mathbb{R}$. Using the triangle inequality, $|x - d'| \leqslant |x - d| + |\Delta|$ where $\Delta = d' - d$, and so it is sufficient to show that

$$\int_A \frac{e^{-\frac{|x-d|}{b}}}{2b} \mathrm{dx} \leqslant e^{\epsilon - \frac{|d'-d|}{b}} \int_A \frac{e^{-\frac{|x-d|}{b}}}{2b} \mathrm{dx} + \delta,$$

for all $d', d \in D, A \in \mathcal{B}(\mathbb{R})$. This last inequality will follow if $1 \leqslant e^{\epsilon - \frac{|\Delta|}{b}} + \delta$ or $b \geqslant \frac{|\Delta|}{\epsilon - \log(1-\delta)}$ for all $\Delta \in \{d' - d : d', d \in D\}$.

Of course, keeping in mind that the l_1 sensitivity of the identity query [10] is precisely given by $\mathrm{diam}(D)$, this example can be seen as an application of the well-known Laplacian mechanism to the identity query.

Example 6. Consider again our earlier example where $D = 2^{\mathcal{H}}$ represents the sets of possible hobbies or interests of people. As noted earlier, it is reasonable to assume that D contains finitely many elements; we denote $|D| = m + 1$. Following Theorem 5 we will construct a mechanism for 1-dimensional databases: this can then be used to define a mechanism for databases in D^n via (7).

For $d \in D$, consider the D-valued random variable Y_d with probability mass function:

$$\mathbb{P}(Y_d = d) = 1 - pm, \quad \mathbb{P}\left(Y_d = d'\right) = p,$$

where $d \neq d' \in D$. We make the reasonable assumption that $1 - pm > p$.

For (ϵ, δ)-differential privacy, we need the following to hold:

$$\mathbb{P}(Y_d \in A) \leqslant e^{\epsilon} \mathbb{P}(Y_{d'} \in A) + \delta, \tag{11}$$

where $A \subseteq D$ and $d, d' \in D$.

We claim that (11) will hold if and only if

$$1 - pm \leqslant e^{\epsilon} p + \delta. \tag{12}$$

This condition is clearly necessary as can be seen by considering the singleton set $A = \{d\}$. To see that it is also sufficient let $d, d' \in D$ and $A \subseteq D$ be given. There are 4 cases to consider.

1. $d, d' \notin A$: Then $\mathbb{P}(Y_d \in A) = \mathbb{P}(Y_{d'} \in A) = p|A|$ and (ϵ, δ)-differential privacy holds trivially.
2. $d, d' \in A$: Then $\mathbb{P}(Y_d \in A) = \mathbb{P}(Y_{d'} \in A) = p(|A| - 1) + 1 - pm = p(|A| - m - 1) + 1$ and (ϵ, δ)-differential privacy holds trivially.
3. $d' \in A, d \notin A$: Then $\mathbb{P}(Y_d \in A) \leqslant \mathbb{P}(Y_{d'} \in A)$ and (ϵ, δ)-differential privacy holds trivially.
4. $d \in A, d' \notin A$: Then

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

From (12), we have

$$1 - pm \leqslant e^{\epsilon} p + \delta = e^{\epsilon}(p|A| - p|A| + p) + \delta \leqslant e^{\epsilon} p|A| - p(|A| - 1) + \delta,$$

since $|A| \geq 1$ ($d \in A$ by hypothesis). Rearranging the above inequality, we see that

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

Thus we can construct an (ϵ, δ)-differentially private mechanism of the form (7) by choosing $p \geqslant \frac{1-\delta}{m+e^{\epsilon}}$.

6. Accuracy

In this section, we consider the question of accuracy for product sanitisations. The literature on the interaction between privacy and accuracy is considerable and previous work has considered examples such as count queries [11], contingency table marginals [1] and spatial data [8]. As product sanitisations are constructed from 1-dimensional mechanisms, we focus on the error of these basic mechanisms here. These results can then be used to derive bounds for data in D^n; the precise form these bounds will take depends on how the metric is constructed on D^n.

We wish to emphasise two points about our work: we are considering a very general class of databases that can incorporate numerical, categorical and functional data; we consider (ϵ, δ)-differential privacy and are not assuming $\delta = 0$.

By the triangle inequality which is valid for arbitrary metrics, $\rho(\cdot, d)$ is a continuous function on D for any fixed d, hence it is measurable with respect to the Borel σ-algebra on D. It follows that $\rho(Y_d, d)$ is a nonnegative-valued random variable. We define the maximal expected error $\mathcal{E}$ as:

$$\mathcal{E} = \max_{d \in D} \mathbb{E}[\rho(Y_d, d)]. \tag{13}$$

For $r > 0$ and $x \in D, B_r(x)$ denotes the open ball

$$B_r(x) := \{y \in D | \rho(y, x) < r\}.$$

We first note that for any differentially private mechanism with $\delta < 1, \mathcal{E} > 0$. If $\delta = 1$, then any mechanism is differentially private.

Lemma 4. *Let a family* $\{Y_d : \Omega \to U | d \in D\}$ *of measurable mappings be given, let* $0 \leqslant \delta < 1$ *and assume that*

$$\mathbb{P}(Y_d \in A) \leqslant e^{\epsilon}\mathbb{P}(Y_{d'} \in A) + \delta, \tag{14}$$

for all $d, d' \in D, A \in \mathcal{A}_U$. *Then* $\mathcal{E} > 0$.

Proof. As D is compact, we can choose u, v in D with $\rho(u, v) = \text{diam}(D)$. Let $r = \frac{\text{diam}(D)}{2}$. Then from (14), it follows that

$$\mathbb{P}(Y_u \in B_r(v)) \geqslant e^{-\epsilon}(\mathbb{P}(Y_v \in B_r(v)) - \delta).$$

As $\rho(x, u) \geqslant r > 0$ for all $x \in B_r(v)$, it follows that $\mathbb{E}[\rho(Y_u, u)] > 0$ unless

$$\mathbb{P}(Y_v \in B_r(v)) = \delta. \tag{15}$$

However, if this is the case then $\mathbb{P}(\rho(Y_v, v) \geqslant r) = 1 - \delta > 0$ and hence $\mathbb{E}[\rho(Y_v, v)] \geqslant r(1 - \delta) > 0$. This completes the proof. □

We now present two simple results giving lower bounds for $\mathcal{E}$. The argument for the following result is inspired by that used to establish Theorem 3.3 of [15].

Theorem 6. *Let a family* $\{Y_d : \Omega \to U | d \in D\}$ *of measurable mappings satisfying* (14) *be given. Then*

$$\mathcal{E} \geqslant (1 - \delta)\left(\frac{\text{diam}(D)}{2(1 + e^{\epsilon})}\right). \tag{16}$$

Proof. Without loss of generality, we may assume that $\mathcal{E}$ is finite. Moreover, from Lemma 4 we know that $\mathcal{E} > 0$. As D is compact, there exist points u, v in D with $\rho(u, v) = \text{diam}(D)$. Set $t = \frac{\text{diam}(D)}{2\mathcal{E}}$; then $t\mathcal{E} = \frac{\text{diam}(D)}{2}$ and the balls $B_{t\mathcal{E}}(u), B_{t\mathcal{E}}(v)$ are disjoint. From Markov's inequality applied to the non-negative random variable $\rho(Y_u, u)$, it follows that

$$\mathbb{P}(Y_u \in B_{t\mathcal{E}}(u)) \geqslant 1 - \frac{1}{t} = 1 - \frac{2\mathcal{E}}{\text{diam}(D)}. \tag{17}$$

It is now immediate that

$$\mathbb{P}(Y_u \in B_{t\mathcal{E}}(v)) \leqslant \frac{2\mathcal{E}}{\text{diam}(D)}. \tag{18}$$

From (14) we know that

$$\mathbb{P}(Y_u \in B_{t\mathcal{E}}(v)) \geqslant e^{-\epsilon}(\mathbb{P}(Y_v \in B_{t\mathcal{E}}(v)) - \delta). \tag{19}$$

Combining (18), (19) and noting that (17) also holds with u replaced by v, we see that

$$\frac{2\mathcal{E}}{\text{diam}(D)} \geqslant e^{-\epsilon}\left(1 - \frac{2\mathcal{E}}{\text{diam}(D)} - \delta\right)$$

and after a simple rearrangement of terms we see that

$$\mathcal{E} \geqslant (1 - \delta)\left(\frac{\text{diam}(D)}{2(1 + e^{\epsilon})}\right),$$

as claimed. □

The previous result applies to any compact metric space D. Now assume that D is discrete so that there exists some $\kappa > 0$ such that

$$\rho(x, y) \geqslant \kappa \quad \forall x, y \in D. \tag{20}$$

This combined with D being compact immediately implies that D is finite. A straightforward alteration of the argument of Theorem 6 yields the following result.

Theorem 7. *Let* D *be finite with* $|D| = m + 1$ *and* $\kappa = \min_{d,d' \in D}\rho(d, d')$. *Let a family* $\{Y_d : \Omega \to U | d \in D\}$ *of measurable mappings be given satisfying* (14). *Then*

$$\mathcal{E} \geqslant (1 - \delta)\left(\frac{\kappa m}{(m + e^{\epsilon})}\right) \tag{21}$$

Proof. It is trivial that the $m+1$ balls $B_{t\mathcal{E}}(u), u \in D$ are all disjoint where $t = \frac{\kappa}{\mathcal{E}}$. Fix some $u \in D$. By the same reasoning as in the proof of Theorem 6,

$$\mathbb{P}(Y_u \in B_{t\mathcal{E}}(u)) \geqslant 1 - \frac{\mathcal{E}}{\kappa}. \tag{22}$$

Moreover, there must exist some $v \neq u$ such that

$$\mathbb{P}(Y_u \in B_{t\mathcal{E}}(v)) \leqslant \frac{\mathcal{E}}{\kappa m}. \tag{23}$$

Choose one such v and apply (14) to obtain

$$\mathbb{P}(Y_u \in B_{t\mathcal{E}}(v)) \geqslant e^{-\epsilon}(\mathbb{P}(Y_v \in B_{t\mathcal{E}}(v)) - \delta). \tag{24}$$

As in the proof of Theorem 6, we can now conclude that

$$\frac{\mathcal{E}}{\kappa m} \geqslant e^{-\epsilon}\left(1 - \frac{\mathcal{E}}{\kappa} - \delta\right).$$

Rearranging this inequality gives us that

$$\mathcal{E} \geqslant (1-\delta)\left(\frac{\kappa m}{m + e^{\epsilon}}\right). \quad \square$$

Example 7. Consider again Example 6. We have shown that there exists an (ϵ, δ)-differentially private mechanism with $p = \frac{1-\delta}{m+e^{\epsilon}}$ where $|D| = m+1$. If D is equipped with the discrete metric so that $\rho(d, d') = 1$ for all $d \neq d'$, then $\kappa = 1$ and for any d, the expected value of $\rho(Y_d, d)$ for this mechanism is given by

$$\sum_{d \neq d'} p = mp = (1-\delta)\left(\frac{m}{m + e^{\epsilon}}\right).$$

So the bound given by Theorem 7 is tight in this simple case.

7. Concluding remarks

We have considered differential privacy in the setting of probability on metric spaces with mechanisms viewed as measurable functions taking values in output query spaces. We have demonstrated that this framework allows mechanisms based on sanitisation and output perturbation to be treated in a uniform manner; moreover we have presented examples to highlight that categorical, functional and numerical data can be treated in this setting. For sanitised response mechanisms, a formal proof that differential privacy with respect to the identity query guarantees differential privacy with respect to any measurable query has been given. We have also introduced the problem of determining sufficient sets for differential privacy, shown that a generating algebra of sets is a sufficient set and applied this fact to functional data in the space $C([0, 1])$ of continuous functions on $[0, 1]$. In the latter half of the paper, we have focussed on product sanitisations of the form (7); we have shown that these mechanisms are (ϵ, δ)-differentially private if and only if the 1-dimensional mechanism used to define them is. This result was then applied in two contexts: to provide a condition for the well-known Laplacian mechanism to be differentially private; and to give a simple example of a differentially private mechanism for discrete categorical data. Finally in Section 6, two simple results giving lower bounds for the maximal expected error for (ϵ, δ) differentially private mechanisms on metric spaces were presented.

Acknowledgments

The first named author was supported by the Science Foundation Ireland Grant SFI/11/PI/1177. The authors would like to thank the anonymous reviewers for their careful reading of an earlier draft of the paper and for their insightful and helpful comments.

References

[1] B. Barak, K. Chaudhuri, C. Dwork, S. Kale, F. McSherry, K. Talwar, Privacy, accuracy, and consistency too: a holistic solution to contingency table release, in: Proceedings of the Twenty-sixth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, 2007, pp. 273–282.
[2] M. Barbaro, T. Zeller Jr. A Face Is Exposed for AOL Searcher No. 4417749, New York Times, August 9, 2006.
[3] A. Beimel, P. Carmi, K. Nissim, E. Weinreb, Private approximation of search problems, SIAM J. Comput. 38 (5) (2008) 1728–1760.
[4] P. Billingsley, Probability and Measure Wiley Series in Probability and Mathematical Statistics, 1995.
[5] D. Boyd, K. Crawford, Six provocations for big data in the Oxford Internet Institutes, in: A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, 2011.

[6] K. Chatzikokolakis, M. Andrés, N.E. Bordenabe, C. Palamidessi, Broadening the scope of differential privacy using metrics, in: Privacy Enhancing Technologies, Springer, Berlin Heidelberg, 2012, pp. 82–102.
[7] Computer Sciences Corporation (CSC), Big Data Universe Beginning to Explode, 2011.
[8] G. Cormode, C. Procopiuc, D. Srivastava, E. Shen, T. Yu, Differentially private spatial decompositions, in: IEEE 28th International Conference on Data Engineering (ICDE), 2012, pp. 20–31.
[9] C. Dwork, Differential privacy, in: Proceedings of the 33rd Annual International Colloquium on Automata, Languages and Programming, LNCS, vol. 4051, Springer, 2006, pp. 1–12.
[10] C. Dwork, Differential privacy: a survey of results, in: 5th International Conference on Theory and Applications of Models of Computation, LNCS, vol. 4978, Springer, 2008, pp. 1–19.
[11] C. Dwork, A firm foundation for private data analysis, Commun. ACM 54 (1) (2011) 86–95.
[12] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, R. Zemel, Fairness through awareness, in: Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, January 2012, pp. 214–226.
[13] A. Ghosh, T. Roughgarden, M. Sundararajan, Universally utility maximising privacy mechanisms, SIAM J. Comput. 41 (6) (2012) 1673–1693.
[14] R. Hall, L. Rinaldo, S. Wasserman, Differential privacy for functions and functional data, J. Mach. Learn. Res. 14 (2013) 703–727.
[15] M. Hardt, K. Talwar, On the geometry of differential privacy, in: Proceedings of the 42nd ACM Symposium on Theory of Computing, 2010, pp. 705–714.
[16] M. Kapralov, K. Talwar, Differentially Private Low Rank Approximation, in: SODA, 2013.
[17] S. Kasiviswanathan, H. Lee, K. Nissim, S. Raskhodnikova, A. Smith, What can we learn privately?, SIAM J Comput. 40 (3) (2011) 793–826.
[18] N. Li, T. Li, S. Venkatasubramanian, t-Closeness: privacy beyond k-anonymity and ℓ-diversity, in: Proceedings of the 21st IEEE International Conference on Data Engineering (ICDE), 2007.
[19] A. Machanavajjhala, D. Kifer, J. Gehrke, M. Venkitasubramaniam, ℓ-Diversity: privacy beyond k-anonymity, ACM Trans. Knowl. Discov. Data 1 (1) (2007) 52 (Article 3).
[20] A. Narayanan, V. Shmatikov, Robust de-anonymization of large sparse datasets, in: 2008 IEEE Symposium on Security and Privacy.
[21] K. Parthasarathy, Probability Measures on Metric Spaces, AMS Chelsea Publishing, 2005 (Reprint).
[22] W. Rudin, Real and Complex Analysis, McGraw-Hill International Editions, 1987.
[23] F. McSherry, I. Mironov, Differentially private recommender systems, in: Proc. KDD, 2009.
[24] J. Soria-Comas, J. Domingo-Ferrer, Optimal data-independent noise for differential privacy, Inform. Sci. 250 (2013) 200–214.
[25] L. Sweenery, Re-identification of De-identified Survey Data. Carnegie Mellon University, School of Computer Science, Data Privacy Laboratory, Technical Report. Pittsburgh: 2000.
[26] L. Sweeney, k-Anonymity: a model for protecting privacy, Int. J. Uncertain. Fuzz. Knowl.-based Syst. 10 (5) (2002) 557–570.
[27] L. Wassermann, S. Zhou, A statistical framework for differential privacy, J. Am. Stat. Assoc. 105 (489) (2010) 375–389.