Differential Privacy: on the trade-off between Utility and Information Leakage

Mário Alvim, Miguel Andrés, Konstantinos Chatzikokolakis, Pierpaolo Degano, Catuscia Palamidessi

Differential Privacy: on the trade-off between Utility and Information Leakage[⋆]

Mário S. Alvim[1], Miguel E. Andrés[1], Konstantinos Chatzikokolakis[1], Pierpaolo Degano[2], and Catuscia Palamidessi[1]

[1] INRIA and LIX, Ecole Polytechnique, France.
[2] Dipartimento di Informatica, Università di Pisa, Italy.

Abstract. Differential privacy is a notion of privacy that has become very popular in the database community. Roughly, the idea is that a randomized query mechanism provides sufficient privacy protection if the ratio between the probabilities that two adjacent datasets give the same answer is bound by e^ϵ. In the field of information flow there is a similar concern for controlling information leakage, i.e. limiting the possibility of inferring the secret information from the observables. In recent years, researchers have proposed to quantify the leakage in terms of min-entropy leakage, a concept strictly related to the Bayes risk. In this paper, we show how to model the query system in terms of an information-theoretic channel, and we compare the notion of differential privacy with that of min-entropy leakage. We show that differential privacy implies a bound on the min-entropy leakage, but not vice-versa. Furthermore, we show that our bound is tight. Then, we consider the utility of the randomization mechanism, which represents how close the randomized answers are to the real ones, in average. We show that the notion of differential privacy implies a bound on utility, also tight, and we propose a method that under certain conditions builds an optimal randomization mechanism, i.e. a mechanism which provides the best utility while guaranteeing ϵ-differential privacy.

1 Introduction

The area of statistical databases has been one of the first communities to consider the issues related to the protection of information. Already some decades ago, Dalenius [12] proposed a famous “ad omnia” privacy desideratum: nothing about an individual should be learnable from the database that could not be learned without access to the database.

1.1 Differential privacy

Dalenius’ property is too strong to be useful in practice: it has been shown by Dwork [13] that no useful database can provide it. In replacement, Dwork

[⋆] This work has been partially supported by the project ANR-09-BLAN-0169-01 PANDA, by the INRIA DRI Equipe Associée PRINTEMPS and by the RAS L.R. 7/2007 project TESLA.

has proposed the notion of *differential privacy*, which has had an extraordinary impact in the community. Intuitively, such notion is based on the idea that the presence or the absence of an individual in the database, or its particular value, should not change in a significant way the probability of obtaining a certain answer for a given query [13–16].

Dwork has also studied sufficient conditions for a randomized function $\mathcal{K}$ to implement a mechanism satisfying ϵ-differential privacy. It suffices to consider a Laplacian distribution with variance depending on ϵ, and mean equal to the correct answer [15]. This is a technique quite diffused in practice.

1.2 Quantitative information flow

The problem of preventing the leakage of secret information has been a pressing concern also in the area of software systems, and has motivated a very active line of research called *secure information flow*. Similarly to the case of privacy, also in this field, at the beginning, the goal was ambitious: to ensure *non-interference*, which means complete lack of leakage. But, as for Dalenius' notion of privacy, non-interference is too strong for being obtained in practice, and the community has started exploring weaker notions. Some of the most popular approaches are the quantitative ones, based on information theory. See for instance [7, 8, 10, 19–21, 24].

The various approaches in the literature mainly differ for the notion of entropy. These notions are related to the kind of attackers we want to model, and to how we measure their success (see [19] for an illuminating discussion of this relation). Most of the approaches are based on Shannon entropy [23], which is used to model an adversary which tries to find out the secret x by asking questions of the form "does x belong to set S?". Shannon entropy is precisely the average number of questions necessary to find out the exact value of x with an optimal strategy (i.e. an optimal choice of the S's). The other most popular notion of entropy in this area is the min-entropy, proposed by Rényi [22]. The corresponding notion of attack is a *single try* of the form "is x equal to value v?". Min-entropy is precisely the logarithm of the probability of guessing the true value with the optimal strategy, which consists, of course, in selecting the v with the highest probability. Approaches based on this notion include [24] and [5].

In this paper, we focus on the approach based on min-entropy. It is worth noting that, while the min-entropy of X, $H_\infty(X)$, represents the a priori probability of success (of the single-try attack), the conditional min-entropy of X given Y, $H_\infty(X \mid Y)$, represents the a posteriori probability of success[1]. This a posteriori probability is the converse of the Bayes risk [11], which has also been used as a measure of the leakage of secret information [4, 6].

[1] We should mention that Rényi did not define the conditional version of the min-entropy, and that there have been various different proposals in literature for this notion. We use here the one proposed by Smith in [24].

1.3 Goal of the paper

The first goal of this paper is to explore the relation between differential privacy and quantitative information flow. We address the problem of characterizing the protection that differential privacy provides us with respect to information leakage. Then, we consider the problem of the utility. This is different from information leakage in that it represents the relation between the reported answer and the true answer. While we want to avoid that the system leaks the information of the participants, we do not need the same protection towards the true answer in itself. It is therefore interesting to explore ways to improve the utility while preserving privacy. We attack this problem by considering the possible structure that the query induces on the true answers.

1.4 Contribution

The main contributions of this paper are the following

- We propose an information-theoretic framework to reason about both information leakage and utility.
- We prove that ϵ-differential privacy implies a bound on the information leakage. The bound is tight.
- We prove that ϵ-differential privacy implies a bound on the utility. We prove that, under certain conditions, the bound is tight.
- We identify a method that, under certain conditions, constructs the randomization mechanisms which maximizes utility while providing ϵ-differential privacy.

1.5 Plan of the paper

The next section introduces some necessary background notions. Section 3 proposes an information-theoretic view of the database query systems, and of its decomposition in terms of the query and of the randomization mechanisms. Section 4 shows that differential privacy implies a bound on the min-entropy leakage, and that the bound is tight. Section 5 shows that differential privacy implies a bound on the utility, and that under certain conditions the bound is tight. Furthermore it shows how to construct and optimal randomization mechanism. Section 6 discusses related work, and Section 7 concludes. The proofs of the results are in the appendix.

2 Background

This section recalls some basic notions on differential privacy and information theory.

2.1 Differential privacy

Roughly, the idea of differential privacy is that a randomized query mechanism provides sufficient privacy protection if the ratio between the probabilities of two different entries to originate a certain answer is bound by e^ϵ, for some given $\epsilon \geq 0$. Dwork's definition of differential privacy is the following:

Definition 1 ([15]). *A randomized function $\mathcal{K}$ satisfies* ϵ-differential privacy *if for all of data sets D' and D'' differing on at most one row, and all $S \subseteq Range(\mathcal{K})$,*

$$Pr[\mathcal{K}(D') \in S] \leq e^\epsilon \times Pr[\mathcal{K}(D'') \in S] \tag{1}$$

2.2 Information theory and interpretation in terms of attacks

In the following, X, Y denote two discrete random variables with carriers $\mathcal{X} = \{x_0, \ldots, x_{n-1}\}$, $\mathcal{Y} = \{y_0, \ldots, y_{m-1}\}$, and probability distributions $p_X(\cdot)$, $p_Y(\cdot)$, respectively. An information-theoretic channel is constituted of an input X, an output Y, and the matrix of conditional probabilities $p_{Y|X}(\cdot \mid \cdot)$, where $p_{Y|X}(y \mid x)$ represent the probability that Y is y given that X is x. We shall omit the subscripts on the probabilities when they are clear from the context.

Min-entropy In [22], Rï¿½nyi introduced a one-parameter family of entropy measures, intended as a generalization of Shannon entropy. The Rï¿½nyi entropy of order α ($\alpha > 0$, $\alpha \neq 1$) of a random variable X is defined as $H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \sum_{x \in \mathcal{X}} p(x)^\alpha$. We are particularly interested in the limit of H_α as α approaches ∞. This is called *min-entropy.* It can be proven that $H_\infty(X) \overset{\text{def}}{=} \lim_{\alpha\to\infty} H_\alpha(X) = -\log_2 \max_{x\in\mathcal{X}} p(x)$.

Rï¿½nyi also defined the α-generalization of other information-theoretic notions, like the Kullback-Leibler divergence. However, he did not define the α-generalization of the conditional entropy, and there is no agreement on what it should be. For the case $\alpha = \infty$, we adopt here the definition of conditional min-entropy proposed by Smith in [24]:

$$H_\infty(X \mid Y) \;=\; -\log_2 \sum_{y\in\mathcal{Y}} p(y) \max_{x\in\mathcal{X}} \; p(x \mid y) \tag{2}$$

Analogously to the Shannon case, we can define the min-entropy leakage I_∞ as $H_\infty(X) - H_\infty(X \mid Y)$, and the capacity C_∞ as $\max_{p_X(\cdot)} I_\infty(X;Y)$. It has been proven in [5] that C_∞ is obtained at the uniform distribution, and that it is equal to the sum of the maxima of each column in the channel matrix, i.e., $C_\infty = \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} p(y \mid x)$.

Interpretation in terms of attacks: Min-entropy can be related to a model of adversary who is allowed to ask exactly one question of the form "is $X = x$?" (one-try attack). More precisely, $H_\infty(X)$ represents the (logarithm of the inverse of the) probability of success for this kind of attacks with the best strategy, which consists, of course, in choosing the x with the maximum probability.

The conditional entropy $H_\infty(X \mid Y)$ represents (the logarithm of) the inverse of (the expected value of) the probability that the same kind of adversary succeeds in guessing the value of X *a posteriori*, i.e. after observing the result of Y. The complement of this probability is also known as *probability of error* or *Bayes risk*. Since in general X and Y are correlated, observing Y increases the probability of success. Indeed we can prove formally that $H_\infty(X \mid Y) \leq H_\infty(X)$, with equality if and only if X and Y are independent. The min-entropy leakage $I_\infty(X;Y) = H_\infty(X) - H_\infty(X|Y)$ corresponds to the *ratio* between the probabilities of success a priori and a posteriori, which is a natural notion of leakage. Note that it is always the case that $I_\infty(X;Y) \geq 0$, which seems desirable for a good notion of leakage.

3 A model of utility and privacy for statistical databases

In this section we present a model of statistical queries on databases, where noise is carefully added to protect privacy and, in general, the reported answer to a query does not need to correspond to the real one. In this model, the notion of information leakage can be used to measure the amount information that an attacker can learn about the database by posting queries and analysing their (reported) answers. Moreover, the model allows us to quantify the utility of the query, that is, how much information about the real answer can be obtained from the reported one. This model will serve as the basis for exploring the relation between differential privacy and information flow.

We fix a finite set $Ind = \{1, 2, \ldots, u\}$ of u individuals participating in the database. In addition, we fix a finite set $Val = \{v_1, v_2, \ldots, v_v\}$, representing the set of (v different) possible values for the *sensitive attribute* of each individual (e.g. disease-name in a medical database)[2]. The absence of an individual in the database, if allowed, can be modeled with one special value in Val (see the discussion at the end of this section). A database $D = \{d_0, \ldots, d_{u-1}\}$ is a u-tuple where each $d_i \in Val$ is the value of the corresponding individual. The set of all databases is $\mathcal{X} = Val^u$. Two databases D, D' are *adjacent*, written $D \sim D'$ iff they differ for the value of exactly one individual.

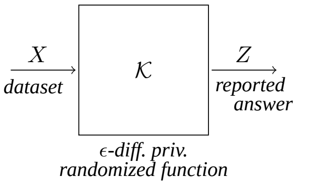


Fig. 1. Randomized function $\mathcal{K}$

Let $\mathcal{K}$ be a randomized function from $\mathcal{X}$ to $\mathcal{Z}$, where $\mathcal{Z} = Range(\mathcal{K})$ (see Figure 1). This function can be modeled by a channel with input and output alphabets $\mathcal{X}, \mathcal{Z}$ respectively. This channel can be specified as usual by a matrix of conditional probabilities $p_{Z|X}(\cdot|\cdot)$. We also denote by X, Z the random variables modeling the input and output of the channel. The definition of differential

[2] In case there are several sensitive attributes in the database (e.g. salary and security number in a census database), we can think of the elements of Val as tuples.

privacy can be directly expressed as a property of the channel: it satisfies ϵ-differential privacy iff

$$p(z|x) \leq e^{\epsilon} p(z|x') \text{ for all } z \in \mathcal{Z}, x, x' \in \mathcal{X} \text{ with } x \sim x'$$

Intuitively, the *correlation* between X and Z measures how much information about the complete database the attacker can obtain by observing the reported answer. We will refer to this correlation as the *leakage* of the channel, denoted by $\mathcal{L}(X,Z)$. In Section 4 we discuss how this leakage can be quantified, using notions from information theory, and we study the behavior of the leakage for differentially private queries.

We then introduce a random variable Y modeling the true answer to the query f, ranging over $\mathcal{Y} = Range(f)$. The correlation between Y and Z measures how much we can learn about the real answer from the reported one. We will refer to this correlation as the *utility* of the channel, denoted by $\mathcal{U}(Y,Z)$. In Section 5 we discuss in detail how utility can be quantified, and we investigate how to construct a randomization mechanism, i.e. a way of adding noise to the query outputs, so that utility is maximized while preserving differential privacy.

In practice, the randomization mechanism is often *oblivious*, meaning that the reported answer Z only depends on the real answer Y and not on the database X. In this case, the randomized function $\mathcal{K}$, seen as channel, can be decomposed into two parts: a channel modeling the query f, and a channel modeling the oblivious randomization mechanism $\mathcal{H}$. The definition of utility in this case is simplified as it only depends on properties of the sub-channel correspondent to $\mathcal{H}$. The leakage relating X and Y and the utility relating Y and Z for a decomposed randomized function are shown in Figure 2.

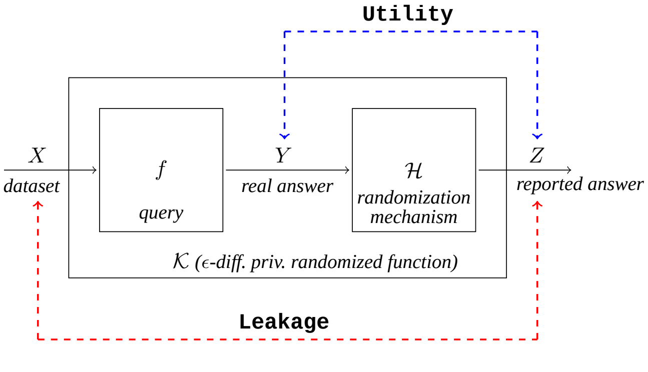


Fig. 2. Leakage and utility for oblivious mechanisms

Leakage about an individual. As already discussed, $\mathcal{L}(X,Z)$ can be used to quantify the information that the attacker can learn about the database as a whole. However, protecting the database as a whole is not the main goal of differential privacy. Indeed, some information will necessarily be revealed, otherwise the query would not be useful. Instead, differential privacy aims at protecting the value of any single individual, even in the worst case where the values of all other individuals are known. To quantify this information leakage we can define smaller channels, where only the information of a specific individual varies. Let $D^- \in \mathit{Val}^{u-1}$ be a $(u-1)$-tuple with the values of all individuals but one (the individual whose degree of protection we want to quantify). We create a channel $\mathcal{K}_{D^-}$ whose input alphabet is the set of all databases in which the $u-1$ other individuals have the same values as in D^-. Intuitively, the information leakage of this channel measures how much information about one particular individual the attacker can learn if the values of all others are known to be D^-. This leakage is studied in Section 4.1.

A note on the choice of values. The choice of the set *Val* depends on the assumptions about the attacker's knowledge. In particular, if the attacker does not know which individuals participate in the database, a distinguished value in *Val* could be interpreted as absence (e.g. 0 or the special value *null*). As discussed in [15], a database D' adjacent to D can be though of either being a superset (or subset) of D with one extra (or missing) row, or being exactly the same database as D in all rows except from one which has a different (non-*null*) value. Our definition of $\sim$ with the possibility of *null* values covers all these cases.

However, an important observation should be made about the choice of *Val*. Most often we are interested in protecting the *actual value* of an individual, not just his participation in the database. In this case, the definition of differential privacy (as well as the channels we are constructing) should include databases with all possible values for each individual, not just the "real" ones. In other words, to prevent the attacker from finding out the individual's value, the probability $p(z|x)$, where x contains the individual's true value, should be close to $p(z|x')$ where x' contains a hypothetical value for this individual.

This might seem unnecessary at first sight, since differential privacy is often though as protecting an participation of an individual in a database. However, hiding the participation of an individual does not imply hiding his value. Consider the following example: we aim at learning the average salary of employees in a small company, and it happens that all of them have exactly the same salary s. We allow anyone to participate or not, while offering ϵ-differential privacy. If we only consider s as the value in all possible databases, then the query is always constant, so answering it any number of times without any noise should satisfy differential privacy for any $\epsilon > 0$. Since all reported answers are s, the attacker can deduce that the salary of all employees, including those not participating in the query, is s. Indeed, the attacker cannot find out who participated, despite the value of all individuals is revealed.

In other cases, we are only interested in hiding the participation (e.g. in a database with information about anonymous donations). Thus, *Val* should be properly selected according to the application. If participation is known and we only wish to hide the values, then *Val* should contain all possible values, e.g., all possible salaries in the example above. If the values are known and participation is to be hidden, then *Val* can contain just the values 0 and 1 denoting absence and presence respectively. Finally, if both the value and the participation are to be protected, then *Val* should contain all values plus *null*.

4 Leakage

As discussed in the previous section, the correlation $\mathcal{L}(X,Z)$ between X and Z measures the information that the attacker can learn about the database by observing the reported answers. In this section, we consider min-entropy leakage as a measure of this information, that is $\mathcal{L}(X,Z) = I_\infty(X;Z)$. We then investigate bounds on information leakage imposed by differential privacy.

Our first result shows that the min-entropy leakage of a randomized function $\mathcal{K}$ is bounded by a quantity depending on ϵ, the numbers u, v of individuals and values respectively. We assume that $v \geq 2$.

Theorem 1. *If $\mathcal{K}$ provides ϵ-*differential privacy *then the min-entropy leakage associated to $\mathcal{K}$ is bounded from above as follows:*

$$I_\infty(X;Z) \leq u \, \log_2 \frac{v\, e^\epsilon}{(v-1+e^\epsilon)}$$

Note that this bound $B(u,v,\epsilon) = u \, \log_2 \frac{v\, e^\epsilon}{(v-1+e^\epsilon)}$ is a continuous function in ϵ, has value 0 when $\epsilon = 0$, and converges to $u \, \log_2 v$ as ϵ approaches infinity. Figure 3 shows the growth of $B(u,v,\epsilon)$ along with ϵ, for various fixed values of u and v.

The following result shows that the bound $B(u,v,\epsilon)$ is *tight*.

Proposition 1. *For every u, v, and ϵ there exists a randomized function $\mathcal{K}$ which provides ϵ-differential privacy and whose min-entropy leakage is $I_\infty(X;Z) = B(u,v,\epsilon)$ for the uniform input distribution.*

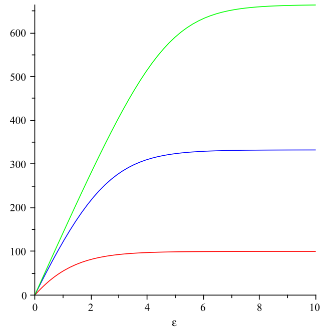
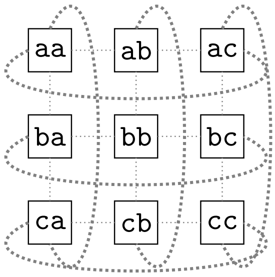


Fig. 3. Graphs of $B(u,v,\epsilon)$ for u=100 and v=2 (lowest line), v=10 (intermediate line), and v=100 (highest line), respectively.

Example 1. Assume that we are interested in the eyes color of a certain population $Ind = \{Alice, Bob\}$. Let $Val = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ where $\mathtt{a}$ stands for *absent* (i.e. the *null* value), $\mathtt{b}$ stands for *blue*, and $\mathtt{c}$ stands for *coal* (black). We can represent each dataset with a tuple $d_1 d_0$, where $d_0 \in Val$ represents the eyes color of *Alice* (cases $d_0 = b$ and $d_0 = c$), or that *Alice* is not in the dataset (case $d_0 = a$). d_1 provides the same kind of information for *Bob*. Note that $v = 3$. Fig 4(a) represents the set $\mathcal{X}$ of all possible datasets and its adjacency relation. Fig 4(b) represents the matrix with input $\mathcal{X}$ which provides ϵ-differential privacy and has the highest min-entropy leakage. In the representation of the matrix, the generic u stands for $\frac{a}{e^{\epsilon\, u}}$, where a is the highest value in the matrix, i.e. $a = \frac{v\, e^\epsilon}{(v-1+e^\epsilon)} = \frac{3\, e^\epsilon}{(2+e^\epsilon)}$.



(a) The datasets and their adjacency relation

	aa	ab	ac	ba	ca	bb	bc	cb	cc
aa	0	1	1	1	1	2	2	2	2
ab	1	0	1	2	2	1	2	1	2
ac	1	1	0	2	2	2	1	2	1
ba	1	2	2	0	1	1	1	2	2
ca	1	2	2	1	0	2	2	1	1
bb	2	1	2	1	2	0	1	1	2
bc	2	2	1	1	2	1	0	2	1
cb	2	1	2	2	1	1	2	0	1
cc	2	2	1	2	1	2	1	1	0

(b) The representation of the matrix

Fig. 4. Universe and highest min-entropy leakage matrix giving ϵ-differential privacy for Example 1.

Note that the bound $B(u, v, \epsilon)$ is guaranteed to be reached with the uniform input distribution. We know from the literature [5, 24] that the I_∞ of a given matrix has its maximum in correspondence of the uniform input distribution, although it may not be the only case.

The construction of the matrix for Proposition 1 gives a square matrix of dimension $v^u \times v^u$. Often, however, the range of $\mathcal{K}$ is fixed, as it is usually related to the possible answers to the query f. Hence it is natural to consider the scenario in which we are given a number $r < v^u$, and want to consider only those $\mathcal{K}$'s whose range has cardinality at most r. Could we, in this restricted setting, find a better bound than the one given by Theorem 1? The following proposition answers this question.

Proposition 2. *Let $\mathcal{K}$ be a randomized function and let $r = |Range(\mathcal{K})|$. If $\mathcal{K}$ provides* ϵ-differential privacy *then the min-entropy leakage associated to $\mathcal{K}$ is bounded from above as follows:*

$$I_\infty(X; Z) \leq \log_2 \frac{r\,(e^\epsilon)^u}{(v - 1 + e^\epsilon)^\ell - (e^\epsilon)^\ell + (e^\epsilon)^u}$$

where $\ell = \lfloor \log_v r \rfloor$.

Note that this bound can be much smaller than the one provided by Theorem 1. For instance, if $r = v$ this bound becomes:

$$\log_2 \frac{v\,(e^\epsilon)^u}{v - 1 + (e^\epsilon)^u}$$

which for large values of u is much smaller than $B(u, v, \epsilon)$.

Let us clarify that there is no contradiction with the fact that the bound $B(u, v, \epsilon)$ is strict: in fact it is strict when we are free to choose the range, but here we fix the dimension of the range.

4.1 Measuring the leakage about an individual

As discussed in Section 3, the main goal of differential privacy is not to protect information about the complete database, but about each individual. To capture the leakage about a certain individual, we start from a tuple $D^{-} \in \mathit{Val}^{u-1}$ containing the given (and known) values of all other $u-1$ individuals. Then we create a channel whose input $X_{D^{-}}$ ranges over all databases where the values of the other individuals are exactly those of D^{-} and only the value of the selected individual varies. Intuitively, $I_{\infty}(X_{D^{-}};Z)$ measures the leakage about the individual's value where all other values are known to be as in D^{-}. As all these databases are adjacent, differential privacy provides a stronger bound for this leakage.

Theorem 2. *If $\mathcal{K}$ provides ϵ-*differential privacy *then for all $D^{-} \in \mathit{Val}^{u-1}$ the min-entropy leakage about an individual is bounded from above as follows:*

$$I_{\infty}(X_{D^{-}};Z) \leq \log_2 e^{\epsilon}$$

Note that this bound is stronger than the one of Theorem 1. In particular, it depends only on ϵ and not on u, v.

5 Utility

As discussed in Section 3, the utility of a randomized function $\mathcal{K}$ is the correlation between the real answers Y for a query and the reported answers Z. In this section we analyze the utility $\mathcal{U}(Y,Z)$ using the classic notion of *utility functions* (see for instance [3]).

For our analysis we assume an oblivious randomization mechanism. As discussed in Section 3, in this case the system can be decomposed into two channels, and the utility becomes a property of the channel associated to the randomization mechanism $\mathcal{H}$ which maps the real answer $y \in \mathcal{Y}$ into a reported answer $z \in \mathcal{Z}$ according to given probability distributions $p_{Z|Y}(\cdot|\cdot)$. However, the user does not necessarily take z as her guess for the real answer, since she can use some Bayesian post-processing to maximize the probability of success, i.e. a right guess. Thus for each reported answer z the user can remap her guess to a value $y' \in \mathcal{Y}$ according to some strategy that maximizes her expected gain. For each pair (y, y'), with $y, y' \in \mathcal{Y}$, there is an associated value given by a gain (or utility) function $g(y, y')$ that represents a score of how useful it is for the user to guess the value y' as the answer when the real answer is y.

It is natural to define the global utility of the mechanism $\mathcal{H}$ as the expected gain:

$$\mathcal{U}(Y,Z) = \sum_{y} p(y) \sum_{y'} p(y'|y) g(y,y') \qquad (3)$$

where $p(y)$ is the prior probability of real answer y, and $p(y'|y)$ is the probability of user guessing y' when the real answer is y.

Assuming that the user uses a remapping function $\rho(z) : \mathcal{Z} \to \mathcal{Y}$, we can derive the following characterization of the utility. We will use δ_x to represent the probability distribution which has value 1 on x and 0 elsewhere.

$$\begin{aligned}
\mathcal{U}(Y,Z) &= \sum_y p(y) \sum_{y'} p(y'|y) g(y,y') && \text{(by (3))} \\
&= \sum_y p(y) \sum_{y'} \left(\sum_z p(z|y) p(y'|z) \right) g(y,y') \\
&= \sum_y p(y) \sum_{y'} \left(\sum_z p(z|y) \delta_{\rho(z)}(y') \right) g(y,y') && \text{(by remap } y' = \rho(z)) \\
&= \sum_y p(y) \sum_z p(z|y) \sum_{y'} \delta_{\rho(z)}(y') g(y,y') \\
&= \sum_{y,z} p(y,z) \sum_{y'} \delta_{\rho(z)}(y') g(y,y') \\
&= \sum_{y,z} p(y,z) g(y, \rho(z))
\end{aligned}$$

A very common utility function is the *binary gain function*, which is defined as $g_{\text{bin}}(y,y') = 1$ if $y = y'$ and $g_{\text{bin}}(y,y') = 0$ if $y \neq y'$. The rationale behind this function is that, when the answer domain does not have a notion of distance, then the wrong answers are all equally bad. Hence the gain is total when we guess the exact answer, and is 0 for all other guesses. Note that if the answer domain is equipped with a notion of distance, then the gain function could take into account the proximity of the reported answer to the real one, the idea being that a close answer, even if wrong, is better than a distant one.

In this paper we do not assume a notion of distance, and we will focus on the binary case. The use of binary utility functions in the context of differential privacy was also investigated in [17][3].

By substituting g with g_{bin} in the above formula we obtain:

$$\mathcal{U}(Y,Z) = \sum_{y,z} p(y,z) \delta_y(\rho(z)) \tag{4}$$

which tells us that the expected utility is the greatest when $\rho(z) = y$ is chosen to maximize $p(y,z)$. Assuming that the user chooses such a maximizing remapping, we have:

$$\mathcal{U}(Y,Z) = \sum_z \max_y p(y,z) \tag{5}$$

This corresponds to the converse of the Bayes risk, and it is closely related to the min conditional entropy and to the min-entropy leakage:

$$H_\infty(Y|Z) = -\log_2 \mathcal{U}(Y,Z) \qquad\qquad I_\infty(Y;Z) = H_\infty(X) + \log_2 \mathcal{U}(Y,Z)$$

[3] The authors of [17] used the dual notion of *loss functions* instead of gain functions, but the final result is equivalent.

5.1 A bound on the utility

In this section we show that the fact that $\mathcal{K}$ provides ϵ-differential privacy induces a bound on the utility. We start by extending the adjacency relation $\sim$ from the datasets $\mathcal{X}$ to the answers $\mathcal{Y}$. Intuitively, the function f associated to the query determines a partition on the set of all databases ($\mathcal{X}$, i.e. Val^u), and we say that two classes are adjacent if they contain an adjacent pair. More formally:

Definition 2. *Given $y, y' \in \mathcal{Y}$, with $y \neq y'$, we say that y and y' are adjacent (notation $y \sim y'$), iff there exist $D, D' \in \mathit{Val}^u$ with $D \sim D'$ such that $y = f(D)$ and $y' = f(D')$.*

Since $\sim$ is symmetric on databases, it is also symmetric on $\mathcal{Y}$, therefore also $(\mathcal{Y}, \sim)$ forms an undirected graph. We define the distance *dist* between two elements $y, y' \in \mathcal{Y}$ as the length of the minimum path from y to y'. For a given natural number d, we use $\mathit{Border}_d(y)$ to denote the set of elements at distance d from y, i.e.

$$\mathit{Border}_d(y) = \{y' \mid \mathit{dist}(y, y') = d\}$$

We recall that a graph automorphism is a permutation of its vertices that preserves its edges. If σ is a permutation of S then an orbit of σ is a set of the form $\{\sigma^i(s) \mid i \in \mathbb{N}\}$ where $s \in S$. A permutation has a single orbit iff $\{\sigma^i(s) | i \in \mathbb{N}\} = S$ for all $s \in S$.

We are now ready to give a bound on the utility:

Theorem 3. *Let $\mathcal{H}$ be a randomization mechanism for the randomized function $\mathcal{K}$ and the query f, and assume that $\mathcal{K}$ provides ϵ-differential privacy. Assume that $(\mathcal{Y}, \sim)$ admits a graph automorphism with only one orbit. Furthermore, assume that there exists a natural number c and an element $y \in \mathcal{Y}$ such that, for every d, either $|\mathit{Border}_d(y)| = 0$ or $|\mathit{Border}_d(y)| \geq c$. Then*

$$\mathcal{U}(X, Y) \leq \frac{(e^\epsilon)^n (1 - e^\epsilon)}{(e^\epsilon)^n (1 - e^\epsilon) + c\,(1 - (e^\epsilon)^n)}$$

where n is the maximum distance from y in $\mathcal{Y}$.

The bound provided by the above theorem is strict, in the sense that, when $|\mathit{Border}_d(y)|$ is exactly c for every d, then we can construct a randomization mechanism $\mathcal{H}$ whose utility achieves the equality in the equation of Theorem 3 and that still provides ϵ-differential privacy. This randomization mechanism is therefore optimal, in the sense that it provides the maximum possible utility for the given ϵ. In the next section we will define formally such an optimal randomization mechanism.

5.2 Constructing an optimal randomization mechanism

Given a query f, and a differential privacy requirement ϵ, it is important to design the randomization mechanism $\mathcal{H}$ in such a way that (together with f) it provides ϵ-differential privacy, but without sacrificing too much on the utility. We show a method to construct the optimal $\mathcal{H}$, at least in some particular cases.

Assume $f : \mathcal{X} \rightarrow \mathcal{Y}$, and consider the graph structure $(\mathcal{Y}, \sim)$ determined by f. Let n be the maximum distance between two nodes in the graph and let c be an integer. We construct the matrix M of conditional probabilities associated to $\mathcal{H}$ as follows. For every column $z \in \mathcal{Z}$ and every row $y \in \mathcal{Y}$, define

$$p_{Z|Y}(z|y) = \frac{\alpha}{(e^\epsilon)^d} \qquad \text{where } d = dist(y, z) \text{ and} \tag{6}$$
$$\alpha = \frac{(e^\epsilon)^n(1 - e^\epsilon)}{(e^\epsilon)^n(1 - e^\epsilon) + c\,(1 - (e^\epsilon)^n)}$$

The following theorem guarantees that the randomization mechanism $\mathcal{H}$ defined above is well defined and optimal, under certain conditions.

Theorem 4. *Let $f : \mathcal{X} \rightarrow \mathcal{Y}$ be a query and let $\epsilon \geq 0$. Assume that $(\mathcal{Y}, \sim)$ admits a graph automorphism with only one orbit, and that there exists c such that, for every $y \in \mathcal{Y}$ and every natural number d, either $|Border_d(y)| = 0$ or $|Border_d(y)| = c$. Then, for such c, the definition in (6) determines a legal channel matrix for $\mathcal{H}$, i.e., for each $y \in \mathcal{Y}$, $p_{Z|Y}(\cdot|y)$ is a probability distribution. Furthermore, the composition $\mathcal{K}$ of f and $\mathcal{H}$ provides ϵ-differential privacy. Finally, $\mathcal{H}$ is optimal in the sense that it maximizes utility when the distribution of Y is uniform.*

The conditions for the construction of the optimal matrix are strong, but there are some interesting scenarios in which they are satisfied. Depending on the degree of connectivity c, we can have $\lfloor\frac{|\mathcal{Y}|}{2}\rfloor - 1$ different cases (note that the case of $c = 1$ is not possible because the datasets are fully connected via their adjacency relation), whose extremes are:

- $(\mathcal{Y}, \sim)$ is a *ring*, i.e. every element has exactly two adjacent elements. This is similar to the case of the counting queries considered in [17], with the difference that our "counting" is in arithmetic modulo $|\mathcal{Y}|$.
- $(\mathcal{Y}, \sim)$ is a *clique*, i.e. every element has exactly $|\mathcal{Y}| - 1$ adjacent elements.

Remark 1. Note that when we have a ring with an even number of nodes the conditions of Theorem 4 are almost met, except that $|Border_d(y)| = 2$ for $d < n$, and $|Border_d(y)| = 1$ for $d = n$, where n is the maximum distance between two nodes in $\mathcal{Y}$. In this case, and if $(e^\epsilon)^2 \geq 2$, we can still construct a legal matrix by doubling the value of such elements. Namely, by defining

$$p_{Z|Y}(z|y) = 2\frac{\alpha}{(e^\epsilon)^n} \qquad \text{if } dist(y, z) = n$$

For all the other elements the definition remains as in (6).

Remark 2. Note that our method can be applied also when the conditions of Theorem 4 are not met: We can always add "artificial" adjacencies to the graph structure so to meet those conditions. Namely, for computing the distance in (6) we use, instead of $(\mathcal{Y}, \sim)$, a structure $(\mathcal{Y}, \sim')$ which satisfies the conditions of Theorem 4, and such that $\sim \subseteq \sim'$. Naturally, the matrix constructed in this way provides ϵ-differential privacy, but in general is not optimal. Of course, the smaller $\sim'$ is, the highest is the utility.

The matrices generated by our algorithm above can be very different, depending on the value of c. The next two examples illustrate queries that give rise to the clique and to the ring structures, and show the corresponding matrices.

Example 2. Consider a database with electoral information where rows corresponds to voters. Let us assume, for simplicity, that each row contains only three fields:

- ID: a unique (anonymized) identifier assigned to each voter;
- CITY: the name of the city where the user voted;
- CANDIDATE: the name of the candidate the user voted for.

Consider the query *"What is the city with the greatest number of votes for a given candidate?"*. For this query the binary function is a natural choice for the gain function: only the right city gives some gain, and any wrong answer is just as bad as any other.

It is easy to see that every two answers are neighbors, i.e. *the graph structure of the answers is a clique.*

Let us consider the scenario where CITY={A,B,C,D,E,F} and assume for simplicity that there is a unique answer for the query, i.e., there are no two cities with exactly the same number of individuals voting for a given candidate. Table 1 shows two alternative mechanisms providing ϵ-differential privacy (with $\epsilon = \log 2$). The first one, M_1, is based on the truncated geometric mechanism method used in [17] for counting queries (here extended to the case where every two answers are neighbors). The second mechanism, M_2, is the one we propose in this paper. Taking the input distribution, i.e. the distribution on Y, as the uniform distribution, it is easy to see that $\mathcal{U}(M_1) = 0.2243 < 0.2857 = \mathcal{U}(M_2)$.

(a) M_1: truncated geometric mechanism

In/Out	A	B	C	D	E	F
A	0.535	0.060	0.052	0.046	0.040	0.267
B	0.465	0.069	0.060	0.053	0.046	0.307
C	0.405	0.060	0.069	0.060	0.053	0.353
D	0.353	0.053	0.060	0.069	0.060	0.405
E	0.307	0.046	0.053	0.060	0.069	0.465
F	0.267	0.040	0.046	0.052	0.060	0.535

(b) M_2: our mechanism

In/Out	A	B	C	D	E	F
A	2/7	1/7	1/7	1/7	1/7	1/7
B	1/7	2/7	1/7	1/7	1/7	1/7
C	1/7	1/7	2/7	1/7	1/7	1/7
D	1/7	1/7	1/7	2/7	1/7	1/7
E	1/7	1/7	1/7	1/7	2/7	1/7
F	1/7	1/7	1/7	1/7	1/7	2/7

Table 1. Mechanisms for the city with higher number of votes for a given candidate

Even for non-uniform distributions, our mechanism still provides better utility. For instance, for $p(A) = p(F) = 1/10$ and $p(B) = p(C) = p(D) = P(E) = 1/5$, we have $\mathcal{U}(M_1) = 0.2412 < 0.2857 = \mathcal{U}(M_2)$. This is not too surprising: the Laplacian method and the geometric mechanism work very well when the domain of answers is provided with a metric and the utility function takes into account the proximity of the reported answer to the real one. It also works well when $(\mathcal{Y}, \sim)$ has low connectivity, in particular in the cases of a ring and of a line. But in this example, we are not in these cases, because we are considering *binary gain functions* and *high connectivity.*

Example 3. Let us consider the same database as the previous example, but now assume a counting query of the form *"What is the number of votes for candidate cand?"*. It is easy to see that each answer has at most two neighbors. More precisely, *the graph structure on the answers is a line.* For illustration purposes, let us assume that only 5 individuals have participated in the election. Table 2 shows two alternative mechanisms providing ϵ-differential privacy ($\epsilon = \log 2$): the truncated geometric mechanism M_1 proposed in [17] and the mechanism we propose M_2, where $c = 2$ and $n = 3$. Note that in order to apply our method we have first to apply Remark 2 to transform the line into a ring, and then Remark 1 to handle the case of the elements at maximal distance from the diagonal.

Le us consider the uniform prior distribution. We see that the utility of M_1 is higher than the utility of M_2, in fact the first is 4/9 and the second is 4/11. This does not contradict our theorem, because our matrix is guaranteed to be optimal only in the case of a ring structure, not a line as we have in this example. If the structure were a ring, i.e. if the last row were adjacent to the first one, then M_1 would not provide ϵ-differential privacy. In case of a line as in this example, the truncated geometric mechanism has been proved optimal [17].

(a) M_1: truncated $\frac{1}{2}$-geom. mechanism

In/Out	0	1	2	3	4	5
0	2/3	1/6	1/12	1/24	1/48	1/48
1	1/3	1/3	1/6	1/12	1/24	1/24
2	1/6	1/6	1/3	1/6	1/12	1/12
3	1/12	1/12	1/6	1/3	1/6	1/6
4	1/24	1/24	1/12	1/6	1/3	1/3
5	1/48	1/48	1/24	1/12	1/6	2/3

(b) M_2: our mechanism

In/Out	0	1	2	3	4	5
0	4/11	2/11	1/11	1/11	1/11	2/11
1	2/11	4/11	2/11	1/11	1/11	1/11
2	1/11	2/11	4/11	2/11	1/11	1/11
3	1/11	1/11	2/11	4/11	2/11	1/11
4	1/11	1/11	1/11	2/11	4/11	2/11
5	2/11	1/11	1/11	1/11	2/11	4/11

Table 2. Mechanisms for the counting query (5 voters)

6 Related work

As far as we know, the first work to investigate the relation between differential privacy and information-theoretic leakage *for an individual* was [1]. In this work, a channel is relative to a given database x, and the channel inputs are all possible databases adjacent to x. Two bounds on leakage were presented, one for the Shannon entropy, and one for the min-entropy. The latter corresponds to Theorem 2 in this paper (note that [1] is an unpublished report).

Barthe and Köpf [2] were the first to investigates the (more challenging) connection between differential privacy and the min-entropy leakage *for the entire universe of possible databases*. They consider only the hiding of the *participation* of individuals in a database, which corresponds to the case of $v = 2$ in our setting. They consider the "end-to-end differentially private mechanisms", which correspond to what we call $\mathcal{K}$ in our paper, and propose, like we do, to interpret them as information-theoretic channels. They provide a bound for the leakage, but point out that it is not tight in general, and show that there cannot be a domain-independent bound, by proving that for any number of individual u the optimal bound must be at least a certain expression $f(u, \epsilon)$. Finally, they show that the question of providing optimal upper bounds for the leakage of $\mathcal{K}$ in terms of rational functions of ϵ is decidable, and leave the actual function as an open question. In our work we used rather different techniques and found (independently) the same function $f(u, \epsilon)$ (the bound $B(u, v, \epsilon)$ in Theorem 1 for $v = 2$), but we proved that $f(u, \epsilon)$ is the optimal bound[4].

Clarkson and Schneider also considered differential privacy as a case study of their proposal for quantification of integrity [9]. There, the authors analyzed database privacy conditions from the literature (such as differential privacy, k-anonymity, and l-diversity) using their framework for utility quantification. In particular, they studied the relationship between differential privacy and a notion of leakage (which is different from ours - in particular their definition is based on Shannon entropy) and they provided a tight bound on leakage.

Heusser and Malacaria [18] were among the first to explore the application of information-theoretic concepts to databases queries. They proposed to model database queries as programs, which allows for statical analysis of the information leaked by the query. However [18] did not attempt to relate information leakage to differential privacy.

In [17] the authors aimed at obtaining optimal-utility randomization mechanisms while preserving differential privacy. The authors proposed adding noise to the output of the query according to the geometric mechanism. Their framework is very interesting because it provides us with a general definition of utility for a randomization mechanism M that captures any possible side information and preference (defined as a loss function) the users of M may have. They proved that the geometric mechanism is optimal in the particular case of counting queries. Our results in Section 5 do not restrict to counting queries, but on the other hand we only consider the case of binary loss function.

[4] When discussing our result with Barthe and Köpf, they said that they also conjectured that $f(u, \epsilon)$ is the optimal bound.

7 Conclusion and future work

An important question in statistical databases is how to deal with the trade-off between the privacy offered to the individuals participating in the database and the utility provided by the answers to the queries. In this work we proposed a model integrating the notions of privacy and utility in the scenario where differential-privacy is applied. We derived a strict bound on the information leakage of a randomized function satisfying ϵ-differential privacy and, in addition, we studied the utility of oblivious differential privacy mechanisms. We provided a way to optimize utility while guaranteeing differential privacy, in the case where a binary gain function is used to measure the utility of the answer to a query.

As future work, we plan to find bounds for more generic gain functions, possibly by using the Kantorovich metric to compare the a priori and a posteriori probability distributions on secrets.

References

1. Mário S. Alvim, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catuscia Palamidessi. Differential privacy versus quantitative information flow. Technical report, 2010.
2. Gilles Barthe and Boris Köpf. Information-theoretic bounds for differentially private mechanisms. In *Proc. of CSF*, 2011. To appear.
3. Jose M. Bernardo and Adrian F. M. Smith. *Bayesian Theory.* J. Wiley & Sons, Inc., 1994.
4. Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Compositional methods for information-hiding. In *Proc. of FOSSACS*, volume 4962 of *LNCS*, pages 443–457. Springer, 2008.
5. Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Quantitative notions of leakage for one-try attacks. In *Proc. of MFPS*, volume 249 of *ENTCS*, pages 75–91. Elsevier, 2009.
6. Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. On the Bayes risk in information-hiding protocols. *J. of Comp. Security*, 16(5):531–571, 2008.
7. David Clark, Sebastian Hunt, and Pasquale Malacaria. Quantitative analysis of the leakage of confidential data. In *Proc. of QAPL*, volume 59 (3) of *Electr. Notes Theor. Comput. Sci*, pages 238–251. Elsevier, 2001.
8. David Clark, Sebastian Hunt, and Pasquale Malacaria. Quantitative information flow, relations and polymorphic types. *J. of Logic and Computation*, 18(2):181–199, 2005.
9. M. R. Clarkson and F. B. Schneider. Quantification of integrity, 2011. Tech. Rep.. `http://hdl.handle.net/1813/22012`.
10. Michael R. Clarkson, Andrew C. Myers, and Fred B. Schneider. Belief in information flow. *J. of Comp. Security*, 17(5):655–701, 2009.
11. Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* J. Wiley & Sons, Inc., second edition, 2006.
12. Tore Dalenius. Towards a methodology for statistical disclosure control. *Statistik Tidskrift*, 15:429 — 444, 1977.

13. Cynthia Dwork. Differential privacy. In *Automata, Languages and Programming, 33rd Int. Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proc., Part II*, volume 4052 of *LNCS*, pages 1–12. Springer, 2006.
14. Cynthia Dwork. Differential privacy in new settings. In *Proc. of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2010, Austin, Texas, USA, January 17-19, 2010*, pages 174–183. SIAM, 2010.
15. Cynthia Dwork. A firm foundation for private data analysis. *Communications of the ACM*, 54(1):86–96, 2011.
16. Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In *Proc. of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 371–380. ACM, 2009.
17. Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In *Proc. of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 351–360. ACM, 2009.
18. Jonathan Heusser and Pasquale Malacaria. Applied quantitative information flow and statistical databases. In *Proc. of the Int. Workshop on Formal Aspects in Security and Trust*, volume 5983 of *LNCS*, pages 96–110. Springer, 2009.
19. Boris Köpf and David A. Basin. An information-theoretic model for adaptive side-channel attacks. In *Proc. of CCS*, pages 286–296. ACM, 2007.
20. Pasquale Malacaria. Assessing security threats of looping constructs. In *Proc. of POPL*, pages 225–235. ACM, 2007.
21. Pasquale Malacaria and Han Chen. Lagrange multipliers and maximum information leakage in different observational models. In *Proc. of PLAS*, pages 135–146. ACM, 2008.
22. Alfréd Rényi. On Measures of Entropy and Information. In *Proc. of the 4th Berkeley Symposium on Mathematics, Statistics, and Probability*, pages 547–561, 1961.
23. Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 625–56, 1948.
24. Geoffrey Smith. On the foundations of quantitative information flow. In *Proc. of FOSSACS*, volume 5504 of *LNCS*, pages 288–302. Springer, 2009.

Appendix

Notation

In the following we assume that A and B are random variables with carriers $\mathcal{A}$ and $\mathcal{B}$, respectively. Let M be a channel matrix with input A and output B. We recall that the matrix M represents the conditional probabilities $p_{B|A}(\cdot|\cdot)$. More precisely, the element of M at the intersection of row $a \in \mathcal{A}$ and column $b \in \mathcal{B}$ is $M_{a,b} = p_{B|A}(b|a)$. Note that if the matrix M and the input random variable A are given, then the output random variable B is completely determined by them, and we use the notation $B(M, A)$ to represent this dependency. We also use $H_\infty^M(A)$ to represent the conditional min-entropy $H_\infty(A|B(M, A))$. Similarly, we use $I_\infty^M(A)$ to denote $I_\infty(A; B(M, A))$.

We denote by $M[l \to k]$ the matrix obtained by "collapsing" the column l into k, i.e.

$$M[l \to k]_{i,j} = \begin{cases} M_{i,k} + M_{i,l} & j = k \\ 0 & j = l \\ M_{i,j} & \text{otherwise} \end{cases}$$

Given a partial function $\rho : \mathcal{A} \to \mathcal{B}$, the image of $\mathcal{A}$ under ρ is $\rho(\mathcal{A}) = \{\rho(a) | a \in \mathcal{A}, \rho(a) \neq \perp\}$, where $\perp$ stands for "undefined".

In the proofs we need to use several indices, hence we typically use the letters i, j, h, k, l to range over rows and columns (usually i, h, l range over rows and j, k range over columns). Given a matrix M, we denote by $\max^j M$ the maximum value of column j over all rows i, i.e. $\max^j M = \max_i M_{i,j}$.

Proofs

For the proofs, it will be useful to consider matrices with certain symmetries. In particular, it will be useful to transform our matrices in square matrices having the property that the elements of the diagonal contain the maximum values of each column, and are all equal. This is the purpose of the following two lemmata: the first one transforms a matrix into a square matrix with all the column maxima in the diagonal, and the second makes all the elements of the diagonal equal. Both transformations preserve ϵ-differential privacy and min-entropy leakage.

Leakage

In this part we prove the results about the bounds on min-entropy leakage. In the following lemmata, we assume that M has input A and output B, and that A has a uniform distribution.

Lemma 1. *Given an $n \times m$ channel matrix M with $n \leq m$, providing ϵ-differential privacy for some $\epsilon \geq 0$, we can construct a square $n \times n$ channel matrix M' such that:*

1. M' provides ϵ-differential privacy.
2. $M'_{i,i} = \max^i M'$ for all $i \in \mathcal{A}$, i.e. the diagonal contains the maximum values of the columns.
3. $H_\infty^{M'}(A) = H_\infty^M(A)$.

Proof. We first show that there exists an $n \times m$ matrix N and an injective total function $\rho : \mathcal{A} \to \mathcal{B}$ such that:

- $N_{i,\rho(i)} = \max^{\rho(i)} N$ for all $i \in \mathcal{A}$,
- $N_{i,j} = 0$ for all $j \in \mathcal{B} \backslash \rho(\mathcal{A})$ and all $i \in \mathcal{A}$.

We iteratively construct ρ, N "column by column" via a sequence of approximating partial functions ρ_s and matrices N_s $(0 \leq s \leq m)$.

- *Initial step* $(s = 0)$.
  Define $\rho_0(i) = \bot$ for all $i \in \mathcal{A}$ and $N_0 = M$.

- s^{th} *step* $(1 \leq s \leq m)$.
  Let j be the s-th column and let $i \in \mathcal{A}$ be one of the rows containing the maximum value of column j in M, i.e. $M_{i,j} = \max^j M$. There are two cases:
  1. $\rho_{s-1}(i) = \bot$: we define
  $$\begin{aligned} \rho_s &= \rho_{s-1} \cup \{i \mapsto j\} \\ N_s &= N_{s-1} \end{aligned}$$
  2. $\rho_{s-1}(i) = k \in \mathcal{B}$: we define
  $$\begin{aligned} \rho_s &= \rho_{s-1} \\ N_s &= N_{s-1}[j \to k] \end{aligned}$$

Since the first step assigns j in ρ_s and the second zeroes the column j in N_s, all unassigned columns $\mathcal{B} \backslash \rho_m(\mathcal{A})$ must be zero in N_m. We finish the construction by taking ρ to be the same as ρ_m after assigning to each unassigned row one of the columns in $\mathcal{B} \setminus \rho_m(\mathcal{A})$ (there are enough such columns since $n \leq m$). We also take $N = N_m$. Note that by construction N is a channel matrix.

Thus we get a matrix N and a function $\rho : \mathcal{A} \to \mathcal{B}$ which, by construction, is injective and satisfies $N_{i,\rho(i)} = \max^{\rho(i)} N$ for all $i \in \mathcal{A}$, and $N_{i,j} = 0$ for all $j \in \mathcal{B} \backslash \rho(\mathcal{A})$ and all $i \in \mathcal{A}$. Furthermore, N provides ϵ-differential privacy because each column is a linear combination of columns of M. It is also easy to see that $\sum_j \max^j N = \sum_j \max^j M$, hence $H_\infty^N(A) = H_\infty^M(A)$ (remember that A has the uniform distribution).

Finally, we create our claimed matrix M' from N as follows: first, we eliminate all columns in $\mathcal{B} \setminus \rho(\mathcal{A})$. Note that all these columns are zero so the resulting matrix is a proper channel matrix, provides differential privacy and has the same conditional min-entropy. Finally, we rearrange the columns according to ρ. Note that the order of the columns is irrelevant, any permutation represents the same conditional probabilities thus the same channel. The resulting matrix M' is $n \times n$ and has all maxima in the diagonal. □

Lemma 2. *Let M be a channel with input and output alphabets $\mathcal{A} = \mathcal{B} = Val^u$, and let $\sim$ be the adjacency relation on Val^u defined in Section 3. Assume that the maximum value of each column is on the diagonal, that is $M_{i,i} = \max^i M$ for all $i \in \mathcal{A}$. If M provides ϵ-differential privacy then we can construct a new channel matrix M' such that:*

1. *M' provides ϵ-differential privacy;*
2. *$M'_{i,i} = M'_{h,h}$ for all $i, h \in \mathcal{A}$ i.e. all the elements of the diagonal are equal;*
3. *$M'_{i,i} = \max^i M'$ for all $i \in \mathcal{A}$;*
4. *$H^M_\infty(A) = H^{M'}_\infty(A)$.*

Proof. Let $k, l \in Val^u$. Recall that $dist(k, l)$ (distance between k and l) is the length of the minimum $\sim$-path connecting k and l, i.e. the number of individuals in which k and l differ. Since $\mathcal{A} = \mathcal{B} = Val^u$ we will use $dist(\cdot,\cdot)$ also between rows and columns. Recall also that $Border_d(h) = \{k \in \mathcal{B} | dist(h, k) = d\}$. For typographical reasons, in this proof we will use the notation $\mathcal{B}_{h,d}$ to represent $Border_d(h)$, and $d(k, l)$ to represent $dist(k, l)$.

Let $n = |\mathcal{A}| = v^u$. The matrix M' is given by

$$M'_{h,k} = \frac{1}{n|\mathcal{B}_{h,d(h,k)}|} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}_{i,d(h,k)}} M_{i,j}$$

We first show that this is a well defined channel matrix, namely $\sum_{k \in \mathcal{B}} M'_{h,k} = 1$ for all $h \in \mathcal{A}$. We have

$$\begin{aligned} \sum_{k \in \mathcal{B}} M'_{h,k} &= \sum_{k \in \mathcal{B}} \frac{1}{n|\mathcal{B}_{h,d(h,k)}|} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}_{i,d(h,k)}} M_{i,j} \\ &= \frac{1}{n} \sum_{i \in \mathcal{A}} \sum_{k \in \mathcal{B}} \frac{1}{|\mathcal{B}_{h,d(h,k)}|} \sum_{j \in \mathcal{B}_{i,d(h,k)}} M_{i,j} \end{aligned}$$

Let $\Delta = \{0, \ldots, u\}$. Note that $\mathcal{B} = \bigcup_{d \in \Delta} \mathcal{B}_{h,d}$, and these sets are disjoint, so the summation over $k \in \mathcal{B}$ can be split as follows

$$\begin{aligned} &= \frac{1}{n} \sum_{i \in \mathcal{A}} \sum_{d \in \Delta} \sum_{k \in \mathcal{B}_{h,d}} \frac{1}{|\mathcal{B}_{h,d}|} \sum_{j \in \mathcal{B}_{i,d}} M_{i,j} \\ &= \frac{1}{n} \sum_{i \in \mathcal{A}} \sum_{d \in \Delta} \sum_{j \in \mathcal{B}_{i,d}} M_{i,j} \sum_{k \in \mathcal{B}_{h,d}} \frac{1}{|\mathcal{B}_{h,d}|} \end{aligned}$$

as $\sum_{k \in \mathcal{B}_{h,d}} \frac{1}{|\mathcal{B}_{h,d}|} = 1$, we obtain

$$= \frac{1}{n} \sum_{i \in \mathcal{A}} \sum_{d \in \Delta} \sum_{j \in \mathcal{B}_{i,d}} M_{i,j}$$

and now the summations over j can be joined together

$$= \frac{1}{n} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}} M_{i,j} = 1$$

We now show that the elements of the diagonal have the intended properties. First, we show that the elements of the diagonal are all the same. We have that $\mathcal{B}_{i,d(h,h)} = \mathcal{B}_{i,0} = \{i\}$ for all $h \in \mathcal{A}$, and therefore:

$$M'_{h,h} = \frac{1}{n} \sum_{i \in \mathcal{A}} M_{i,i}$$

Then, we show that they are the maxima for each column. Note that $|\mathcal{B}_{i,d}| = \binom{u}{d}(v-1)^d$ which is independent from i. We have:

$$\begin{aligned}
M'_{h,k} &= \frac{1}{n|\mathcal{B}_{h,d(h,k)}|} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}_{i,d(h,k)}} M_{i,j} \\
&\leq \frac{1}{n|\mathcal{B}_{h,d(h,k)}|} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}_{i,d(h,k)}} M_{i,i} \qquad (M \text{ has maxima in the diag.}) \\
&= \frac{1}{n} \sum_{i \in \mathcal{A}} \frac{|\mathcal{B}_{i,d(h,k)}|}{|\mathcal{B}_{h,d(h,k)}|} M_{i,i} \\
&= \frac{1}{n} \sum_{i \in \mathcal{A}} M_{i,i} = M'_{h,h}
\end{aligned}$$

It easily follows that $\sum_j \max^j M' = \sum_j \max^j M$ which implies that $H_\infty^M(A) = H_\infty^{M'}(A)$.

It remains to show that M' provides ϵ-differential privacy, namely that

$$M'_{h,k} \leq e^\epsilon M'_{h',k} \qquad \forall h, h', k \in \mathcal{A} : h \sim h'$$

Since $d(h, h') = 1$, by the triangular inequality we derive:

$$d(h', k) - 1 \leq d(h, k) \leq d(h', k) + 1$$

Thus, there are exactly 3 possible cases:

1. $d(h, k) = d(h', k)$.
   The result is immediate since $M'_{h,k} = M'_{h',k}$.

2. $d(h, k) = d(h', k) - 1$.
   Define
   $$\mathcal{S}_{i,j} = \{j' \in \mathcal{B}_{i,d(i,j)+1} | j' \sim j\}$$

Note that $|\mathcal{S}_{i,j}| = (u - d(i,j))(v-1)$ (i and j are equal in $u - d(i,j)$ elements, and we can change any of them in $v-1$ ways). The following holds:

$$\begin{aligned}
M_{i,j} &\le e^\epsilon M_{i,j'} \qquad \forall j' \in \mathcal{S}_{i,j} && \text{(diff. privacy)} \Rightarrow \\
(u - d(i,j))(v-1) M_{i,j} &\le e^\epsilon \sum_{j' \in \mathcal{S}_{i,j}} M_{i,j'} && \text{(sum of the above)} \Rightarrow \\
\sum_{j \in \mathcal{B}_{i,d(h,k)}} (u - d(h,k))(v-1) M_{i,j} &\le e^\epsilon \sum_{j \in \mathcal{B}_{i,d(h,k)}} \sum_{j' \in \mathcal{S}_{i,j}} M_{i,j'} && \text{(sum over } j\text{)}
\end{aligned}$$

Let $d = d(h,k)$. Note that each $j' \in \mathcal{B}_{i,d+1}$ is contained in exactly $d+1$ different sets $\mathcal{S}_{i,j}, j \in \mathcal{B}_{i,d}$. So the right-hand side above sums all elements of $\mathcal{B}_{i,d+1}$, $d+1$ times each. Thus we get

$$(u-d)(v-1) \sum_{j \in \mathcal{B}_{i,d}} M_{i,j} \le e^\epsilon \; (d+1) \sum_{j \in \mathcal{B}_{i,d+1}} M_{i,j} \tag{7}$$

Finally, we have

$$\begin{aligned}
M'_{h,k} &= \frac{1}{n|\mathcal{B}_{h,d}|} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}_{i,d}} M_{i,j} \\
&\le e^\epsilon \; \frac{1}{n\binom{u}{d}(v-1)^d} \frac{d+1}{(u-d)(v-1)} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}_{i,d+1}} M_{i,j} \qquad \text{(from (7))} \\
&\le e^\epsilon \; \frac{1}{n\binom{u}{d+1}(v-1)^{d+1}} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{B}_{i,d+1}} M_{i,j} \\
&= e^\epsilon M'_{h',k}
\end{aligned}$$

3. $d(h,k) = d(h',k) + 1$.
   Symmetrical to the the case $d(h,k) = d(h',k) - 1$.

□

We are now ready to prove our first main result.

Theorem 1. If $\mathcal{K}$ provides *ϵ-differential privacy* then the min-entropy leakage associated to $\mathcal{K}$ is bounded from above as follows:

$$I_\infty(X;Z) \le u \, \log_2 \frac{v \, e^\epsilon}{(v - 1 + e^\epsilon)}$$

Proof. Let us assume, without loss of generality, that $|\mathcal{X}| \le |\mathcal{Z}|$ (if this is not the case, then we add enough zero columns, i.e. columns containing only 0's, so to match the number of rows. Note that adding zero columns does not change the min-entropy leakage).

For our proof we need a square matrix with all column maxima on the diagonal, and all equal. We obtain such a matrix by transforming the matrix associated to $\mathcal{K}$ as follows: first we apply Lemma 1 to it (with $A = X$ and $B = Z$), and then we apply Lemma 2 to the result of Lemma 1. The final matrix M has size $n \times n$, with $n = |\mathcal{X}| = v^u$, provides ϵ-differential privacy, and for all rows i, h we have that $M_{i,i} = M_{h,h}$ and $M_{i,i} = \max^i M$. Furthermore, $I_\infty^M(X)$ is equal to the min-entropy leakage of $\mathcal{K}$.

Let us denote by α the value of every element in the diagonal of M, i.e. $\alpha = M_{i,i}$ for every row i. Note that for every $j \in \mathit{Border}_d(i)$ (i.e. every j at distance d from a given i) the value of $M_{i,j}$ can be at most $\frac{M_{i,i}}{(e^\epsilon)^d}$, hence $M_{i,j} \leq \frac{\alpha}{(e^\epsilon)^d}$. Furthermore each element j at distance d from i can be obtained by changing the value of d individuals in the u-tuple representing i. We can choose those d individuals in $\binom{u}{d}$ possible ways, and for each of these individuals we can change the value (with respect to the one in i) in $v-1$ possible ways. Therefore $|\mathit{Border}_d(i)| = \binom{u}{d}(v-1)^d$, and we obtain:

$$\sum_{d=0}^{u} \binom{u}{d} (v-1)^d \frac{\alpha}{(e^\epsilon)^d} \;\leq\; \sum_{j=1}^{n} M_{i,j}$$

Since each row represents a probability distribution, the elements of row i must sum up to 1. Hence:

$$\sum_{d=0}^{u} \binom{u}{d} (v-1)^d \frac{\alpha}{(e^\epsilon)^d} \;\leq\; 1$$

Now we apply some simple transformations:

$$\begin{aligned}
\sum_{d=0}^{u} \binom{u}{d} (v-1)^d \frac{\alpha}{(e^\epsilon)^d} \;&\leq\; 1 \qquad\qquad \Longleftrightarrow \\
\sum_{d=0}^{u} \binom{u}{d} (v-1)^d \frac{\alpha}{(e^\epsilon)^d} \;&\leq\; 1 \qquad\qquad \Longleftrightarrow \\
\alpha \sum_{d=0}^{u} \binom{u}{d} (v-1)^d ((e^\epsilon)^d)^{u-d} \;&\leq\; (e^\epsilon)^u
\end{aligned}$$

Since $\alpha \sum_{d=0}^{u} \binom{u}{d} (v-1)^d (e^\epsilon)^{u-d} \;=\; (v-1+e^\epsilon)^u$ (binomial expansion), we obtain:

$$\alpha \leq \left(\frac{e^\epsilon}{v-1+e^\epsilon} \right)^u \tag{8}$$

Therefore:

$$\begin{aligned}
I_\infty^M(X) &= H_\infty(X) - H_\infty^M(X) && \text{(by definition)}\\
&= \log_2 v^u + \log_2 \sum_j \alpha \frac{1}{n}\\
&= \log_2 v^u + \log_2 \alpha\\
&\leq \log_2 v^u + \log_2 \left(\frac{e^\epsilon}{v-1+e^\epsilon}\right)^u && \text{(by (8))}\\
&= u \log_2 \frac{v\, e^\epsilon}{v-1+e^\epsilon}
\end{aligned}$$

□

Next proposition shows that the bound obtained in previous theorem is tight.

Proposition 1. For every u, v, and ϵ there exists a randomized function $\mathcal{K}$ which provides ϵ-differential privacy and whose min-entropy leakage, for the uniform input distribution, is $I_\infty(X;Z) = B(u,v,\epsilon)$.

Proof. The adjacency relation in $\mathcal{X}$ determines a graph structure $G_\mathcal{X}$. Set $\mathcal{Z} = \mathcal{X}$ and define the matrix of $\mathcal{K}$ as follows:

$$p_\mathcal{K}(z|x) = \frac{B(u,v,\epsilon)}{(e^\epsilon)^d} \qquad \text{where } d \text{ is the distance between } x \text{ and } z \text{ in } G_\mathcal{X}$$

It is easy to see that $p_\mathcal{K}(\cdot|x)$ is a probability distribution for every x, that $\mathcal{K}$ provides ϵ-differential privacy, and that $I_\infty(X;Z) = B(u,v,\epsilon)$. □

We consider now the case in which $|Range(\mathcal{K})|$ is bounded by a number smaller than v^u.

In the following when we have a random variable X, and a matrix M with row indices in $\mathcal{A} \subsetneq \mathcal{X}$, we will use the notations $H_\infty^M(X)$ and $I_\infty^M(X)$ to represent the conditional min-entropy and leakage obtained by adding "dummy raws" to M, namely rows that extend the input domain of the corresponding channel so to match the input X, but which do not contribute to the computation of $H_\infty^{M'}(X)$. Note that it is easy to extend M this way: we only have to make sure that for each column j the value of each of these new rows is dominated by $\max^j M'$.

We will also use the notation $\sim_u$ and $\sim_\ell$ to refer to the standard adjacency relations on Val^u and Val^ℓ, respectively.

Lemma 3. *Let $\mathcal{K}$ be a randomized function with input X, where $\mathcal{X} = Val^{Ind}$, providing ϵ-differential privacy. Asssume that $r = |Range(\mathcal{K})| = v^\ell$, for some $\ell < u$. Let M be the matrix associated to $\mathcal{K}$. Then it is possible to build a square matrix M' of size $v^\ell \times v^\ell$, with row and column indices in $\mathcal{A} \subseteq \mathcal{X}$, and a binary relation $\sim' \subseteq \mathcal{A} \times \mathcal{A}$ such that $(\mathcal{A}, \sim')$ is isomorphic to $(Val^\ell, \sim_\ell)$, and such that:*

1. *$M'_{i,j} \leq (e^{\epsilon})^{u-l+d}\, M'_{i,j}$ for all $i, j, k \in \mathcal{A}$, where d is the $\sim'$-distance between j and k.*
2. *$M'_{i,i} = M'_{h,h}$ for all $i, h \in \mathcal{A}$, i.e. elements of the diagonal are all equal*
3. *$M'_{i,i} = \max^i M'$ for all $i \in \mathcal{A}$, i.e. the diagonal contains the maximum values of the columns.*
4. *$H^{M'}_{\infty}(X) = H^{M}_{\infty}(X)$.*

Proof. We first apply a procedure similar to that of Lemma 1 to construct a square matrix of size $v^\ell \times v^\ell$ which has the maximum values of each column in the diagonal. (In this case we construct an injection from the columns to rows containing their maximum value, and we eliminate the rows that at the end are not associated to any column.) Then define $\sim'$ as the projection of $\sim_u$ on Val^ℓ. It is easy to see that point 1 in this lemma is satisfied by this definition of $\sim'$. Finally, apply the procedure in Lemma 2 (on the structure $(\mathcal{A}, \sim')$) to make all elements in the diagonal equal. Note that this procedure preserves the property in point 1, and conditional min-entropy. Hence $H^{M'}_{\infty}(X) = H^{M}_{\infty}(X)$. □

Proposition 2. Let $\mathcal{K}$ be a randomized function and let $r = |\mathit{Range}(\mathcal{K})|$. If $\mathcal{K}$ provides *ϵ-differential privacy* then the min-entropy leakage associated to $\mathcal{K}$ is bounded from above as follows:

$$I_{\infty}(X; Z) \;\leq\; \log_2 \frac{r\,(e^{\epsilon})^u}{(v-1+e^{\epsilon})^{\ell} - (e^{\epsilon})^{\ell} + (e^{\epsilon})^u}$$

where $\ell = \lfloor \log_v r \rfloor$.

Proof. Assume first that r is of the form v^ℓ. We transform the matrix M associated to $\mathcal{K}$ by applying Lemma 3, and let M' be the resulting matrix. Let us denote by α the value of every element in the diagonal of M', i.e. $\alpha = M'_{i,i}$ for every row i, and let us denote by $\mathit{Border'}_d(i)$ the set of elements whose $\sim'$-distance from i is d. Note that for every $j \in \mathit{Border'}_d(i)$ we have that $M'_{i,i} \leq M'_{i,j}(e^{\epsilon})^{u-\ell+d}$, hence

$$M'_{i,j} \leq \frac{\alpha}{(e^{\epsilon})^{u-\ell+d}}$$

Furthermore each element j at $\sim'$-distance d from i can be obtained by changing the value of d individuals in the ℓ-tuple representing i (remember that $(\mathcal{A}, \sim')$ is isomorphic to $(\mathit{Val}^\ell, \mathit{sim}_\ell)$). We can choose those d individuals in $\binom{\ell}{d}$ possible ways, and for each of these individuals we can change the value (with respect to the one in i) in $v-1$ possible ways. Therefore

$$|\mathit{Border'}_d(i)| = \binom{\ell}{d}(v-1)^d$$

Taking into account that for $M'_{i,i}$ we do not need to divide by $(e^{\epsilon})^{u-\ell+d}$, we obtain:

$$\alpha + \sum_{d=1}^{\ell} \binom{\ell}{d}(v-1)^d \frac{\alpha}{(e^{\epsilon})^{u-\ell+d}} \;\leq\; \sum_j M'_{i,j}$$

Since each row represents a probability distribution, the elements of row i must sum up to 1. Hence:

$$\alpha + \sum_{d=1}^{u} \binom{u}{d} (v-1)^d \frac{\alpha}{(e^\epsilon)^{u-\ell+d}} \;\leq\; 1 \tag{9}$$

By performing some simple calculations, similar to those of the proof of Theorem 1, we obtain:

$$\alpha \;\leq\; \frac{(e^\epsilon)^u}{(v-1+e^\epsilon)^\ell - (e^\epsilon)^\ell + (e^\epsilon)^u}$$

Therefore:

$$\begin{aligned}
I_\infty^{M'}(X) \;&=\; H_\infty(X) - H_\infty^{M'}(X) && \text{(by definition)} \\
&=\; \log_2 v^u + \log_2 \sum_{j=1}^{v^\ell} \alpha \frac{1}{v^u} \\
&=\; \log_2 v^u + \log_2 \frac{1}{v^u} + \log_2 (v^\ell\, \alpha) \\
&\leq\; \log_2 \frac{v^\ell\, (e^\epsilon)^u}{(v-1+e^\epsilon)^\ell - (e^\epsilon)^\ell + (e^\epsilon)^u} && \text{(by (9))}
\end{aligned} \tag{10}$$

Consider now the case in which r is not of the form v^ℓ. Let ℓ be the maximum integer such that $v^\ell < r$, and let $m = r - v^\ell$. We transform the matrix M associated to $\mathcal{K}$ by collapsing the m columns with the smallest maxima into the m columns with highest maxima. Namely, let $j_1, j_2, \ldots, j_m$ the indices of the columns which have smallest maxima values, i.e. $\max^{j_t} M \leq \max^j M$ for every column $j \neq j_1, j_2, \ldots, j_m$. Similarly, let $k_1, k_2, \ldots, k_m$ be the indexes of the columns which have maxima values. Then, define

$$N = M[j_1 \to k_1][j_2 \to k_2] \ldots [j_m \to k_m]$$

Finally, eliminate the m zero-ed columns to obtain a matrix with exactly v^ℓ columns. It is easy to show that

$$I_\infty^M(X) \;\leq\; I_\infty^N(X) \frac{r}{v^\ell}$$

After transforming N into a matrix M' with the same min-entropy leakage as described in the first part of this proof, from (10) we conclude

$$I_\infty^M(X) \;\leq\; I_\infty^{M'}(X) \frac{r}{v^\ell} \;\leq\; \log_2 \frac{r\, (e^\epsilon)^u}{(v-1+e^\epsilon)^\ell - (e^\epsilon)^\ell + (e^\epsilon)^u}$$

□

We now turn our attention to the min-entropy leakage associated to an individual.

Lemma 4. *If a randomized function $\mathcal{K} : A \to B$ respects an ϵ-ratio in the sense that $p_{\mathcal{K}}(b|a') \leq e^{\epsilon} \cdot p_{\mathcal{K}}(b|a'')$ for all $a', a'' \in \mathcal{A}$ and $b \in \mathcal{B}$, then the min-entropy leakage from A to B is bounded by:*

$$I_\infty(A; B) \leq \epsilon \log_2 e$$

Proof. For clarity reasons, in this proof we use the notation $p(b|A = a)$ for the probability distributions $p_{\mathcal{K}}(b|A = a)$ associated to $\mathcal{K}$.

$$\begin{aligned}
-H_\infty(A|B) &= \log_2 \textstyle\sum_b p(b) \max_a p(a|b) && \text{(by definition)}\\
&= \log_2 \textstyle\sum_b \max_a (p(b)\, p(a|b)) && \\
&= \log_2 \textstyle\sum_b \max_a (p(a)\, p(b|a)) && \text{(by the Bayes theorem)}\\
&\leq \log_2 \textstyle\sum_b \max_a (p(a)\, e^{\epsilon}\, p(b|\hat{a})) && \text{(by hypothesis on } \mathcal{K}\text{, for some fixed } \hat{a})\\
&= \log_2 \textstyle\sum_b e^{\epsilon}\, p(b|\hat{a}) \max_a p(a) && \\
&= \log_2 (e^{\epsilon} \max_a p(a) \textstyle\sum_b p(b|\hat{a})) && \\
&= \log_2 (e^{\epsilon} \max_a p(a)) && \text{(by probability laws)}\\
&= \log_2 e^{\epsilon} + \log \max_a p(a) && \\
&= \epsilon \log_2 e - H_\infty(A) && \text{(by definition)}
\end{aligned}$$

Therefore:

$$H_\infty(A|B) \geq H_{\infty(A)} - \epsilon \log_2 e \qquad (11)$$

This gives us a bound on the min-entropy leakage:

$$\begin{aligned}
I_\infty(A; B) &= H_\infty(A) - H_\infty(A|B)\\
&\leq \epsilon \log_2 e \qquad\qquad \text{(by (11))}
\end{aligned}$$

□

Theorem 2. If $\mathcal{K}$ provides ϵ-*differential privacy* then for all $D^- \in \mathit{Val}^{u-1}$ the min-entropy leakage about an individual is bounded from above as follows:

$$I_\infty(X_{D^-}; Z) \leq \log_2 e^{\epsilon}$$

Proof. By construction, the elements of $\mathcal{X}_{D^-}$ are all adjacent. Hence $\mathcal{K}_{D^-}$ respects an ϵ-ratio. Thus we are allowed to apply Lemma 4 (with $X = X_{D^-}$ and $\mathcal{K} = \mathcal{K}_{D^-}$), which gives immediately the intended result. □

Utility

In this part we prove the results on utility. We start with a lemma which plays a role analogous to Lemma 2, but for a different kind of graph structure: in this case, we require the graph to have an automorphism with a single orbit.

Lemma 5. *Let M be the matrix of a channel with the same input and output alphabet $\mathcal{A}$. Assume an adjacency relation $\sim$ on $\mathcal{A}$ such that the graph $(\mathcal{A}, \sim)$ has an automorphism σ with a single orbit. Assume that the maximum value of each column is on the diagonal, that is $M_{i,i} = \max^i M$ for all $i \in \mathcal{A}$. If M provides ϵ-differential privacy then we can construct a new channel matrix M' such that:*

1. *M' provides ϵ-differential privacy;*
2. *$M'_{i,i} = M'_{h,h}$ for all $i, h \in \mathcal{A}$;*
3. *$M'_{i,i} = \max^i M'$ for all $i \in \mathcal{A}$;*
4. *$H^M_\infty(A) = H^{M'}_\infty(A)$.*

Proof. Let $n = |\mathcal{A}|$. For every $h, k \in \mathcal{A}$ let us define the elements of M' as:

$$M'_{h,k} = \frac{1}{n} \sum_{i=0}^{n-1} M_{\sigma^i(h),\sigma^i(k)}$$

First we prove that M' provides ϵ-differential privacy. For every pair $h \sim l$ and every k:

$$\begin{aligned} M'_{h,k} &= \sum_{i=0}^{n-1} M_{\sigma^i(h),\sigma^i(k)} \\ &\leq \sum_{i=0}^{n-1} e^\epsilon M_{\sigma^i(l),\sigma^i(k)} && \text{(by ϵ-diff. privacy, for some l s.t. $\rho(\sigma^i(h')) = k$)} \\ &= e^\epsilon M'_{l,k} \end{aligned}$$

Now we prove that for every h, $M'_{h,\cdot}$ is a legal probability distribution. Remember that $\{\sigma^i(k) | 0 \leq i \leq n-1\} = \mathcal{A}$ since σ has only one orbit.

$$\begin{aligned} \sum_{k=0}^{n-1} M'_{h,k} &= \sum_{k=0}^{n-1} \frac{1}{n} \sum_{i=0}^{n-1} M_{\sigma^i(h),\sigma^i(k)}, \\ &= \sum_{i=0}^{n-1} \frac{1}{n} \sum_{k=0}^{n-1} M_{\sigma^i(h),\sigma^i(k)} \\ &= \sum_{i=0}^{n-1} \frac{1}{n} 1 && \text{(since $\{\sigma^i(k) | 0 \leq i \leq n-1\} = \mathcal{A}$)} \\ &= 1 \end{aligned}$$

Next we prove that the diagonal contain the maximum values of the columns, i.e., for every k, $M'_{k,k} = \max^k M'$.

$$\begin{aligned}
M'_{k,k} &= \frac{1}{n}\sum_{k=0}^{n-1} M_{\sigma^i(k),\sigma^i(k)} \\
&\geq \frac{1}{n}\sum_{k=0}^{n-1} M_{\sigma^i(h),\sigma^i(k)} \quad (\text{since } M_{\sigma^i(h),\sigma^i(h)} = \max^{\sigma^i(h)} M) \\
&= M'_{hk}
\end{aligned}$$

Finally, we prove that $I_\infty^{M'}(A) = I_\infty^M(A)$. It is enough to prove that $H_\infty^{M'}(A) = H_\infty^M(A)$.

$$\begin{aligned}
H_\infty^{M'}(A) &= \sum_{h=0}^{n-1} M_{h,h} \\
&= \frac{1}{n}\sum_{h=0}^{n-1}\sum_{i=0}^{n-1} M_{\sigma^i(h),\sigma^i(h)} \quad (\text{since } \{\sigma^i(h) | 0 \leq i \leq n-1\} = \mathcal{A}) \\
&= \frac{1}{n}\sum_{h=0}^{n-1} H_\infty^M(A) \quad (\text{since } M_{\sigma^i(h),\sigma^i(h)} = \max^{\sigma^i(h)} M) \\
&= H_\infty^M(A)
\end{aligned}$$

□

Theorem 3. Let $\mathcal{H}$ be a randomization mechanism for the randomized function $\mathcal{K}$ and the query f, and assume that $\mathcal{K}$ provides ϵ-differential privacy. Assume that $(\mathcal{Y}, \sim)$ admits a graph automorphism with only one orbit. Furthermore, assume that there exists a natural number c and an element $y \in \mathcal{Y}$ such that, for every d, either $|Border_d(y)| = 0$ or $|Border_d(y)| \geq c$. Then

$$\mathcal{U}(X,Y) \leq \frac{(e^\epsilon)^n(1-e^\epsilon)}{(e^\epsilon)^n(1-e^\epsilon) + c\,(1-(e^\epsilon)^n)}$$

where n is the maximum distance from y in $\mathcal{Y}$.

Proof. Consider the matrix M obtained by applying Lemma 1 to the matrix of $\mathcal{H}$, and then Lemma 5 to the result of Lemma 1. Let us call α the value of the elements in the diagonal of M.

Let us take an element $M_{i,i} = \alpha$. For each element $j \in Border_d(M_{i,i})$, the value of $M_{i,j}$ can be at most $\frac{\alpha}{e^{d\epsilon}}$. Also, the elements of row i represent a probability distribution, so they sum up to 1. Hence we obtain:

$$\alpha + \sum_{d=1}^{n} |Border_d(y)| \frac{\alpha}{(e^\epsilon)^d} \;\leq\; 1$$

Now we perform some simple calculations:

$$\begin{aligned}
\alpha + \sum_{d=1}^{n} |Border_d(y)| \frac{\alpha}{(e^\epsilon)^d} \;&\leq\; 1 \implies \text{(since by hypothesis } |Border(y,d)| \geq c\text{)}\\
\alpha + \sum_{d=1}^{n} c\, \frac{\alpha}{(e^\epsilon)^d} \;&\leq\; 1 \iff\\
\alpha\,(e^\epsilon)^n + c\,\alpha \sum_{d=1}^{n} (e^\epsilon)^{n-d} &\leq (e^\epsilon)^n \iff\\
\alpha\,(e^\epsilon)^n + c\,\alpha \sum_{t=0}^{n-1} (e^\epsilon)^{t} &\leq (e^\epsilon)^n \iff \text{(geometric progression sum)}\\
\alpha\,(e^\epsilon)^n + c\,\alpha\,\frac{1-(e^\epsilon)^n}{1-e^\epsilon} &\leq (e^\epsilon)^n \iff\\
\alpha \leq \frac{(e^\epsilon)^n(1-e^\epsilon)}{(e^\epsilon)^n(1-e^\epsilon) + c\,(1-(e^\epsilon)^n)}
\end{aligned}$$

Since $\mathcal{U}(Y,Z) = \alpha$, we conclude.

□

Theorem 4. Let $f : \mathcal{X} \to \mathcal{Y}$ be a query and let $\epsilon \geq 0$. Assume that $(\mathcal{Y}, \sim)$ admits a graph automorphism with only one orbit, and that there exists c such that, for every $y \in \mathcal{Y}$ and every natural number d, either $|Border_d(y)| = 0$ or $|Border_d(y)| = c$. Then, for such c, the definition in (6) determines a legal channel matrix for $\mathcal{H}$, i.e., for each $y \in \mathcal{Y}$, $p_{Z|Y}(\cdot|y)$ is a probability distribution. Furthermore, the composition $\mathcal{K}$ of f and $\mathcal{H}$ provides ϵ-differential privacy. Finally, $\mathcal{H}$ is optimal in the sense that it maximizes utility when the distribution of Y is uniform.

Proof. We follow a reasoning analogous to the proof of Theorem 3, but using $|Border(y,d)| = c$, to prove that

$$\mathcal{U}(Y,Z) = \frac{(e^\epsilon)^n(1-e^\epsilon)}{(e^\epsilon)^n(1-e^\epsilon) + c\,(1-(e^\epsilon)^n)}$$

From the same theorem, we know that this is a maximum for the utility. □