Project no.:	**PIRSES-GA-2011-295261**
Project full title:	**Mobility between Europe and Argentina applying Logics to Systems**
Project Acronym:	**MEALS**
Deliverable no.:	**3.2 / 1**
Title of Deliverable:	**On the information leakage of differentially-private mechanisms**

Contractual Date of Delivery to the CEC:	**1-Apr-2013**
Actual Date of Delivery to the CEC:	**20-Mar-2013**
Organisation name of lead contractor for this deliverable:	**INR**
Author(s):	**Mário S. Alvim, Miguel E. Andrés, Konstantinos Chatzikokolakis, Pierpaolo Degano, Catuscia Palamidessi**
Participants(s):	**INR, UNC, TUE, UBA**
Work package contributing to the deliverable:	**WP3**
Nature:	**R**
Dissemination Level:	**Public**
Total number of pages:	**46**
Start date of project:	1 Oct. 2011 **Duration:** 48 month

Abstract:

The concept of *differential privacy* emerged as an approach to protect the privacy of the individuals participating in statistical databases. Roughly, a mechanism satisfies differential privacy if the presence or value of a single individual in a database does not change significantly the likelihood of obtaining a certain answer to any statistical query posed by the data analyst. Differentially-private mechanisms are often *oblivious*: first the query is processed on the database to produce the true answer, and then this answer is adequately randomized before being reported to the data analyst. Ideally the mechanism should minimize leakage, i.e., obfuscate as much as possible the link between the reported answer and the individuals' data. At the same time, it should maximize utility, i.e., the reported answer should be as close as possible to the true one. These two goals are, however, conflicting, and a trade-off between privacy and utility is imposed.

In this paper we use quantitative information flow to analyze the leakage and the utility of oblivious differentially-private mechanisms. We introduce a technique that exploits some graph-symmetries presented by the adjacency relation on databases to derive bounds on the leakage of the mechanism, measured as min-entropy leakage. We use *identity gain functions*, which are closely related to min-entropy leakage, to evaluate utility, and therefore we are also able to derive bounds for it. Depending on the graph-symmetries we consider, we can additionally build a mechanism that maximizes utility while preserving the required level of differential privacy.

This project has received funding from the European Union Seventh Framework Programme (FP7 2007-2013) under Grant Agreement Nr. 295261.

Contents

1 Introduction

In statistical databases the data of a (large) number of individuals are collected, and data analysts are allowed to pose statistical queries. Typical examples of statistical queries include average value, total counting, or the fraction of the entries in the database that satisfy some property. Statistical databases are of special importance in many areas. For instance, medical databases can provide information on how a disease spreads, and a census database can help authorities decide how to spend the budget of years to come. The field of *statistical disclosure control* concerns the problem of revealing accurate statistics about a set of respondents while preserving the privacy of individuals.

In principle we would like to consider the *global information* relative to the database as *public*, and the *individual information* about a individual as *private*. It is not trivial to combine these two requirements, because these two kinds of information are intrinsically linked. As an example, suppose that a statistical database of individuals of a certain country indicates that the life expectancy for women is 5 years longer than for men. Clearly this piece of information reveals something about each individual in this country, *even about individuals not present in the database*.

Moreover, private information about particular individuals can be revealed even if only statistical queries about the sample as a whole are allowed. Consider, for example, a database that stores the values of the salaries of a set of individuals, and assume that a data analyst can pose the query "what is the average salary of the individuals in the database?". We would like to obtain this information without being able to infer the salary of any specific individual. Unfortunately this is not always possible. In particular, if the number of individuals in the database is known (via the result of a counting query, for instance), and an individual is removed from (or included in) the database, it is possible to infer his salary by querying again the database and calculating the influence of the removal (or inclusion) on the answer to the query.

Yet another issue to be considered is *side information*, which is any piece of data about individuals that the adversary has and that does not come from the database itself. It may originate from prior beliefs, from public sources such as newspapers, or even from other databases. The problem with side information is that some piece of information in itself harmless may pose a threat to the privacy of individuals if combined with the answer to a query in the database [11].

The problem of statistical disclosure control is not trivial, and in order to cope with it, Dwork proposed the notion of *differential privacy*, which has received a lot of attention in the privacy community. Differential privacy is based on the idea that the presence or absence of an individual in the database, or the individual's particular value, should not significantly change the probability of obtaining any specific answer for a certain query [11, 12, 13, 14]. Intuitively, it ensures that individuals can opt in (or out) of the database without significantly changing the probability of any given answer to a query to be reported, and hence it is "safe" for an individual to join (or leave) the database. A strong property of differential privacy is that it is *indifferent to side information*: the privacy guarantee holds no matter what side information an adversary may have.

There are several approaches in the literature to implement differentially-private query mechanisms. For numeric queries, Dwork showed that it is enough to consider additive *Laplacian* noise [11] that calculates the true answer to the query on the database, and then adds to this answer some noise drawn from a Laplace distribution. Later on more sophisticated mechanisms have been proposed, which correlate noise between queries and allow for more information to be extracted from the database while still preserving differential privacy [7, 22, 16]. For the case where output perturbation does not make sense (e.g., the case of non-numeric queries, or the case where the perturbed answer is not almost as good as the exact answer), McSherry and Talwar proposed the *exponential mechanism* [21].

Although differential privacy is a promising approach to the problem of statistical disclosure control, the fact that it relies on the randomization of the query answer poses some challenges with respect to the *utility* of the mechanism. If the noise is not introduced with sufficient care, the reported answer can be so "different" from the true answer that the informative purpose of the database is compromised.

Indeed, a differentially-private mechanism has two goals. On one hand, in order to ensure privacy, it needs to minimize the amount of information that the randomized answers reveal about the database (and, in particular, about individuals). Drawing a parallel with information flow, privacy can be seen as the converse of the *information leakage* from databases to reported answers. On the other hand, the mechanism also needs to provide good *utility*, i.e., it needs to maximize the correlation between the true answers and the randomized answers. Clearly differential privacy introduces a trade-off between privacy and utility: the noise that obfuscates the link between the reported answer and the database has the side-effect of driving the answer away from the true one.

In this context, some natural questions arise. How to quantify the privacy and the utility a differentially-private mechanism provides? Is there a way of ensuring a desired level of differential privacy and providing the maximum utility at the same time? In this paper we address these two questions by connecting differential privacy to the well-established field of *quantitative information flow*.

1.1 Quantitative information flow

Protecting private information is a concern not only of the privacy community. In a broad range of computational systems it is necessary to avoid the leakage of secret information through public observables. If some information is supposed to be confidential, then unauthorized users should not be allowed to infer such information from the output or the behavior of the system. Ideally we would like systems to be completely secure, i.e., protect the secret information entirely, but this goal is usually impossible to achieve. It is thus important to quantify leakage, and that is the concern of the field of *quantitative information flow*.

Information theory is widely regarded as a natural framework to provide firm foundations to quantitative information flow. The basic idea is that a system can be seen as a channel in the information-theoretic sense, where the secret is the input and the observables are the output. The entropy of the input provides a measure of its *vulnerability*, i.e., how easily an adversary can discover the secret. As argued in [20], the notion of entropy should be chosen according to the model of adversary, and to the way we estimate the success of the attack. The entropy of a random variable represents its *uncertainty*, hence the vulnerability is anti-monotonic on the entropy. One of the main notions of entropy used for quantitative information flow in the literature is *min-entropy* [23, 8, 24].

Independently of the intended model of adversary, the notion of leakage can be expressed in a uniform way as the difference between the *initial* uncertainty about the secret, i.e., the uncertainty *before* we run the system, and the *remaining* uncertainty about the secret, i.e., the uncertainty *after* we run the system and observe its outcome:

$$\textit{information leakage} = \textit{initial uncertainty} - \textit{remaining uncertainty} \tag{1}$$

In general the observation of the outcome should increase the probabilistic knowledge about the secret, and consequently decrease the corresponding uncertainty. Therefore we expect the result of (1) to be non-negative. This is indeed the case for min-entropy.

1.2 Goal of this paper

The main goal of this work is investigate the relationships between differential privacy and quantitative information flow. At the motivational level, the concern about privacy is akin the concern about information leakage. At the conceptual level, the differentially-private mechanism can be seen as an information-theoretic channel, and the limit case of complete privacy ($\epsilon = 0$) corresponds to a 0-capacity channel, which does not allow any leakage.

We use these similarities to describe and measure relevant properties of the differentially-private mechanism in terms of uncertainty, entropy, and leakage. More specifically, we investigate the notion of differential privacy and its implications in the light of the min-entropy framework for information flow.

First, we address the problem of characterizing the protection that differential privacy grants with respect to information leakage. Given that a trade-off between privacy and utility is an imposition on any differentially-private query mechanism, it is interesting to quantify the utility of the mechanism and explore ways to improve it while preserving privacy. We attack this problem by considering the graph structure that the query induces on the true answers.

Contribution The main contributions of this paper can be summarized as follows.

- We propose an information-theoretic framework to reason about both information leakage and utility in differentially-private query mechanisms.
- We explore the graph-theoretic foundations of the adjacency relation on databases, and we point out two types of symmetries (distance-regular graphs and vertex-transitive graphs) which allows us to prove that differential privacy induces a bound on the min-entropy leakage.
- Furthermore, we prove that this bound is tight, i.e., that there always exists a differentially private mechanism that attains this bound.

- We prove that if the graph structure of the answers satisfies our symmetry conditions, then differential privacy induces a bound on the utility, measured in terms of identity gain functions. We also prove that this bound is tight.
- As a side result, we prove that under our symmetry conditions the exponential mechanism is optimal, i.e., it provides maximum utility for a given degree of privacy.

1.3 Roadmap to the paper

This paper is organized as follows. In Section 2 we review some basic concepts of information theory, quantitative information flow, differential privacy, and graph theory.

In Section 3 we introduce a model to reason about privacy and utility for query mechanisms for statistical databases. We focus on oblivious mechanisms, i.e., randomization methods that can be split into two sub-channels in cascade. The first channel corresponds to the query: it takes a database as input and produces the true answer as output. The second channel corresponds to the noise: it takes the true answer as input, and outputs a randomized answer (according to some differentially-private policy) to be reported to the data analyst. In this model, the leakage of private information is the information that flows from databases to reported answers, i.e., the information flow in the composition of the two sub-channels. We measure this leakage using the min-entropy leakage framework. Utility is corresponds to posterior min-entropy in the second sub-channel.

To derive bounds on leakage and utility, we develop a technique that exploits the graph structure induced by the adjacency relation on databases and on true answers. In Section 4 we show that if the graph structure of the input to a channel is either distance-regular or vertex-transitive, it is possible to transform the channel matrix while preserving their information flow properties. This allows us to calculate bounds on the a posterior entropy of the channels, and to derive bounds on the maximum min-entropy leakage of the channels.

In Section 5 we show that the set of all databases are both distance-regular and vertex-transitive, and we derive our results for leakage by applying our bounds on min-entropy leakage to the channel from databases to reported answers. In Section 6 we consider the adjacency relation on true answers induced by the adjacency relation on databases, and then apply our bounds to the channel from true answers to reported answers. This allows us to derive our results for utility.

Finally, in Section 7 we review some of the related work in the literature, and in Section 8 we make our final remarks and conclude the paper.

2 Preliminaries

In this section we review some concepts from from differential privacy, from the information-theoretic framework for quantitative information flow, and from graph theory.

2.1 The information-theoretic framework for quantitative information flow

In the following, let 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 let π denote a probability distribution for X, called a *prior distribution*.

An *information-theoretic channel* (or simply a *channel*) is a triple $(\mathcal{X}, \mathcal{Y}, M)$, where $\mathcal{X}$ is the *channel's input*, $\mathcal{Y}$ is the *channel's output*, and M is a matrix of conditional probabilities called the *channel matrix*. Each element $M_{x,y}$ represents the probability that Y takes value y given that X has value x. Together, π and M, define induce a joint probability distribution p for X, Y, defined as $p(x, y) = \pi_x M_{x,y}$. Note that p satisfies $p(x) = \pi_x$ and $p(y|x) = M_{x,y}$.

There are several measures of the information shared by X and Y via the channel matrix. In this paper we will concentrate in the measures based on min-entropy leakage.

Min-entropy leakage The *vulnerability* of X is defined as

$$V(X) = \max_{x \in \mathcal{X}} p(x)$$

Intuitively, $V(X)$ represents the probability that an adversary can correctly guess the value of X in a single try, since a rational adversary will choose as her guess the value of x with maximum probability. Correspondingly, the *conditional vulnerability* of X given Y is defined as

$$\begin{aligned} V(X|Y) &= \sum_{y \in \mathcal{Y}} p(y) V(X|Y = y) \\ &= \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} p(x) p(y|x) \end{aligned}$$

In words, it represents the probability that the adversary can correctly guess the value of X in one try *after* observing the value of Y. It can be shown that $\frac{V(X|Y)}{V(X)} \geq 1$, and intuitively it represents by how much the adversary's probability of success is increased by the observation of the channel's output.

For mathematical convenience, the vulnerability is usually converted into bits by taking its logarithm in base 2. The *min-entropy* of X is then defined as

$$H_\infty(X) = -\log_2 V(X)$$

and the *conditional min-entropy* of X given Y is defined as

$$H_\infty(X|Y) = -\log_2 V(X|Y)$$

The *min-entropy leakage* (or simply *min-leakage*) of X to Y is defined as the difference between the *a priori* (before observing the value of Y) and *a posteriori* (after observing the value of Y) min-entropies:

$$\begin{aligned} I_\infty(X; Y) &= H_\infty(X) - H_\infty(X|Y) \\ &= \log_2 \frac{V(X|Y)}{V(X)} \end{aligned}$$

It can be shown that $0 \leq I_\infty(X;Y) \leq H_\infty(X)$, but min-leakage is not symmetric, i.e., in general we have $I_\infty(X;Y) \neq I_\infty(Y;X)$.

Min-capacity is the worst-case leakage obtained by maximising over all input distributions:

$$C_\infty = \max_{\pi} I_\infty(X;Y)$$

It has been proven in [8] 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., $C_\infty = \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} p(y|x)$.

2.2 Differential Privacy

Let $\mathcal{X}$ be the set of all possible databases. Two databases $x, x' \in \mathcal{X}$ are *adjacent* (or *neighbors*), written $x \sim x'$, if they differ for the presence or the value of exactly one individual. We call $\sim$ the *adjacency relation* on databases. Note that the structure $(\mathcal{X}, \sim)$ forms an undirected graph, where vertices are databases and edges connect every two neighbor databases.

A differentially-private mechanism (or simply a mechanism) is a probabilistic function from $\mathcal{X}$ to some set of possible answers $\mathcal{Z}$. Differential privacy is based on the idea that a mechanism provides sufficient protection if the ratio between the probabilities of two adjacent databases to give a certain answer is bound by e^ϵ, for some $\epsilon > 0$. Formally:

Definition 1 ([13]). A mechanism $\mathcal{K}$ from $\mathcal{X}$ to $\mathcal{Z}$ satisfies ϵ-differential privacy if for all pairs $x, x' \in \mathcal{X}$, with $x \sim x'$, and all $S \subseteq \mathcal{Z}$, we have:

$$Pr[\mathcal{K}(x) \in S] \leq e^\epsilon \times Pr[\mathcal{K}(x') \in S] \tag{2}$$

Note that in this work we consider $\mathcal{X}, \mathcal{Z}$ to be finite, therefore all probability distributions are discrete and it is sufficient to consider probabilities of the form $Pr[\mathcal{K}(x) = z]$ in the above definition.

Intuitively, (2) implies that if a value of a single individual changes in a dataset (either by inclusion, removal or modification), the probability of the querying mechanism to report a specific answer will not “vary much”. In other words, the influence of a single individual in a database is “negligible” with respect to the whole set of individuals. Of course the notion of what is meant by “much” and “negligible” depends on the value of ϵ, and it is usual to think of ϵ as a constant smaller than 1. In fact, for $0 < \epsilon < 1$ we have $e^\epsilon \approx (1 + \epsilon)$, which further attests the notion that differential privacy ensures a “small” multiplicative difference in the distributions generated by neighbor databases.

2.3 Graph theory

Given a graph $G = (\mathcal{V}, \sim)$, the *distance* $d(v, w)$ between two vertices $v, w \in \mathcal{V}$ is the number of edges in a shortest path connecting them. If $\mathcal{V}$ is the set of vertices of G, we denote by $\mathcal{V}_{\langle d \rangle}(v)$ the subset of vertices in $\mathcal{V}$ that are at distance d from the vertex v. The *diameter* δ of G is the maximum distance between any two vertices in $\mathcal{V}$, i.e., $\delta = \max_{v,w \in \mathcal{V}} d(v, w)$. The *degree* of a vertex is the number of edges incident to it. G is called *regular* if every vertex has the same

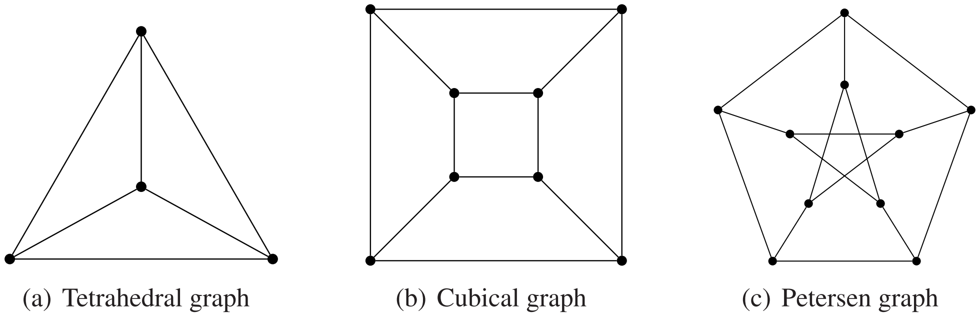

(a) Tetrahedral graph (b) Cubical graph (c) Petersen graph

Figure 1: Some distance-regular graphs with degree 3

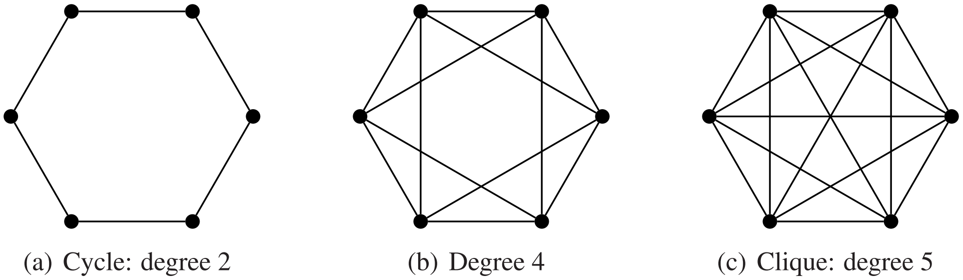

(a) Cycle: degree 2 (b) Degree 4 (c) Clique: degree 5

Figure 2: Some vertex-transitive graphs

degree. A regular graph with vertices of degree k is called a *k-regular graph*. An *automorphism* of G is a permutation σ on the vertex set $\mathcal{V}$ such that for any pair of vertices v, w, if $v \sim w$, then $\sigma(v) \sim \sigma(w)$. The *automorphism group* of a graph G is the set of all of its automorphisms. Let σ^k be the composition of σ with itself k times, i.e., $\sigma^k(v) = \sigma(\sigma^{k-1}(v))$ for $k > 0$, and $\sigma^0(v) = v$.

The following two definitions introduce the classes of graphs that we will be interested in, and some examples of distance-regular graphs and vertex-transitive graphs can be found in Figure 1 and Figure 2, respectively.

Definition 2 (Distance-regular graph)**.** A graph $G = (\mathcal{V}, \sim)$ is called *distance-regular* if there exist integers b_d and c_d ($d \in \{0, \ldots, \delta\}$) (called *intersection numbers*) such that, for all vertices v, w at distance $d(v, w) = d$, there are exactly

- b_d neighbors of w in $\mathcal{V}_{\langle d+1 \rangle}(v)$
- c_d neighbors of w in $\mathcal{V}_{\langle d-1 \rangle}(v)$

The second class we are interested in are vertex-transitive graphs.

Definition 3 (Vertex-transitive graph)**.** A graph $G = (\mathcal{V}, \sim)$ is called *vertex-transitive* if for any pair $v, w \in \mathcal{V}$ there exists an automorphism σ such that $\sigma(v) = w$.

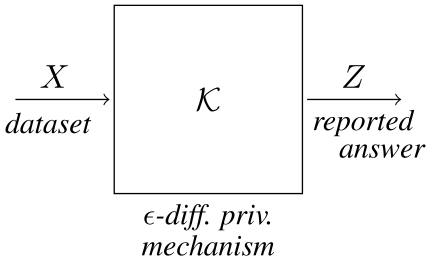


Figure 3: Mechanism $\mathcal{K}$

3 A model of utility and privacy for statistical databases

We fix a finite set $\mathcal{U} = \{0, 1, \ldots, u-1\}$ of u individuals participating in the database. In addition, we fix a finite set $\mathcal{V} = \{v_0, v_1, \ldots, v_{v-1}\}$, often called the *universe*, representing the set of (v different) possible values for the *sensitive attribute* of each individual (e.g., disease name in a medical database). In the more general case where there are several sensitive attributes (e.g., salary and security number in a census sample), we can think of the elements of $\mathcal{V}$ as tuples. The absence of an individual in the database, if allowed, can be modeled with a special value in $\mathcal{V}$. A database $x = d_0 \ldots d_{u-1}$ is a u-tuple where each $d_i \in \mathcal{V}$ is the value of the corresponding individual. The set of all databases is $\mathcal{X} = \mathcal{V}^u$. Denoting by $\sim$ the *adjacency relation* on databases (i.e., $x \sim x'$, if and only if x and x' are databases that differ in the value of exactly one individual), we recall that the structure $(\mathcal{X}, \sim)$ is an undirected graph.

Let $\mathcal{K}$ be a mechanism from $\mathcal{X}$ to $\mathcal{Z}$ (see Figure 3). The mechanism can be modeled by a channel $(\mathcal{X}, \mathcal{Z}, M)$, where $\mathcal{X}, \mathcal{Z}$ are the channel's inputs and outputs respectively, and M is the channel matrix. The definition of differential privacy can be directly expressed as a property of the channel: it satisfies ϵ-differential privacy iff:

$$M_{x,z} \leq e^{\epsilon} M_{x,z'} \qquad \text{for all } x, x' \in \mathcal{X} \text{ s.t. } x \sim x', \text{ and all } z \in \mathcal{Z}$$

Intuitively, the correlation between X and Z measures how much information about the complete database the adversary can obtain by observing the reported answer. We will refer to this correlation as the *leakage* of the channel, denoted by $\mathcal{L}(X, Z)$.

We model the true answer to the query f by the random variable Y ranging over $\mathcal{Y} = Range(f)$. The correlation between Y and Z measures how much we can learn about the true answer from the reported one. We will refer to this correlation as the *utility* of the channel, denoted by $\mathcal{U}(Y, Z)$.

In practice, the mechanism $\mathcal{K}$ is often *oblivious*, meaning that the reported answer Z only depends on the true answer Y and not on the database X. In this case, $\mathcal{K}$, seen as a channel, can be decomposed into two parts: a channel from X to Y modelling the query f, and a noise channel $\mathcal{H}$ from Y to Z. These two channels are said to be *in cascade*, since the output of the first one is the input for the second one. In this case, the definition of utility only depends on the noise channel $\mathcal{H}$. Figure 4. depicts the leakage relating X and Y and the utility relating Y and Z for a decomposed oblivious mechanism.

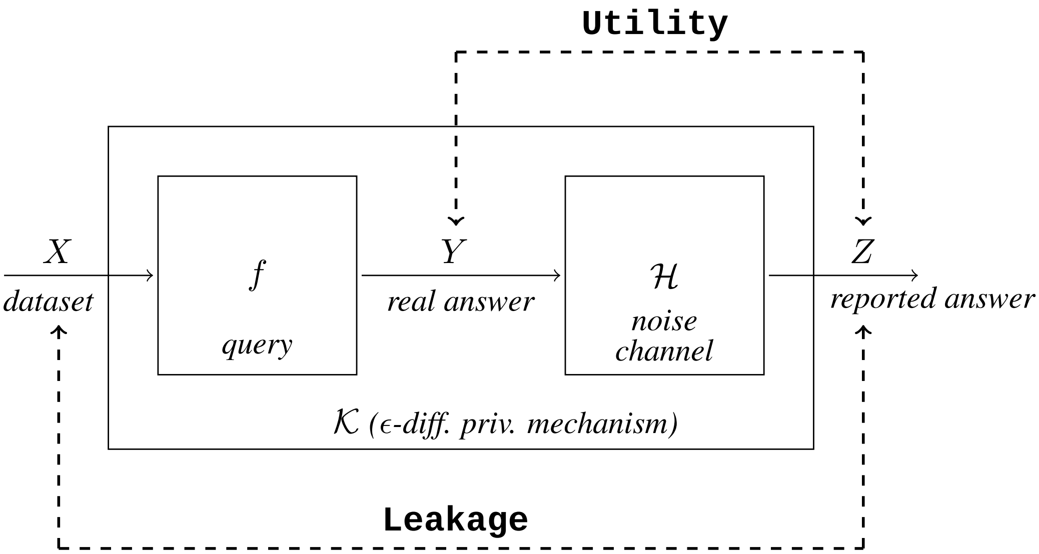


Figure 4: Leakage and utility for oblivious mechanisms

We capture the notion of the adversary's side information as the prior distribution on X, which is standard in works on information flow and also on differential privacy [15, 19].

The main goal of this paper is to explore the relation between differential privacy and quantitative information flow. We will use the model we just introduced to explore the following questions:

1. Does ϵ-differential privacy induce a bound on the information leakage of the mechansim $\mathcal{K}$?
2. Does ϵ-differential privacy induce a bound on the utility?
3. Given a query f and a value $\epsilon > 0$, can we construct a mechanism $\mathcal{K}$ which satisfies ϵ-differential privacy and also presents maximum utility?

We will see that the answer to question 1 is positive in case we take the measure of leakage to be the min-entropy leakage, and we provide bounds that are tight (i.e., for every ϵ there is a $\mathcal{K}$ whose leakage reaches the bound). Regarding question 2, we are able to give a tight bound in some cases which depend on the structure of the query, and for the same cases, we are able to construct an oblivious $\mathcal{K}$ with maximum utility (defined in terms of a identity gain function), as requested in question 3.

4 Relating differential privacy and quantitative information flow

In this section we propose a general technique that exploits the graph structure of the input to channels in order to establish relations between ϵ-differential privacy and min-entropy leakage, which imply results regarding privacy and utility. More specifically, if the graph structure of the input to a channel is distance-regular or vertex-transitive, we show how to transform the channel matrix into an equivalent matrix with certain regularities that can be used to derive bounds on the a posteriori min-entropy of the channel. These bounds can be instantiated for our results on the leakage of mechanisms (Section 5) and their utility (Section 6).

In Section 4.2 we will present our transformations on the channel matrix, and in Section 4.3 we will show how to derive bounds on the posterior min-entropy for the matrix obtained. It is important to note that the bounds on the posterior min-entropy are obtained under the assumption that the input distribution is uniform. This is not a restriction for our bounds on the leakage: as seen in Section 2, the maximum min-entropy leakage is achieved in the uniform input distribution and, therefore, any bound for the uniform distribution is also a bound for all other input distributions. In the case of utility the assumption of uniform input distribution is more restrictive, but we will see that it still provides interesting results for several practical cases.

4.1 Assumptions and notation

In the rest of this section we consider channels (usually referred to by M, M', M'' or N) having input A and output B, with finite carriers $\mathcal{A} = \{a_0, \ldots, a_{n-1}\}$ and $\mathcal{B} = \{b_0, \ldots, b_{m-1}\}$, respectively, and we assume that the probability distribution of A is uniform. Furthermore, we assume that $|\mathcal{A}| = n \leq |\mathcal{B}| = m$. If it is the case that $n > m$, we just add to the matrix enough zero-ed columns, i.e., columns containing only 0's, so as to match the number of rows. Note that adding zero-ed columns does not change the min-entropy leakage nor the conditional min-entropy of the channel. We assume as well an adjacency relation $\sim$ on $\mathcal{A}$, i.e., that $(\mathcal{A}, \sim)$ is an undirected graph structure. With a slight abuse of notation, we may use the number i to denote the element a_i of $\mathcal{A}$ (or, equivalently, the element b_i of $\mathcal{B}$) whenever it is clear from the context. In particular, we also write $i \sim h$ when i and h are associated with adjacent elements $a_i \sim a_h$ of $\mathcal{A}$, and we write $d(i, h)$ to denote the distance between the elements of $\mathcal{A}$ associated with i and h.

We recall that a channel matrix M satisfies ϵ-differential privacy if for each column j and for each pair of rows i and h such that $i \sim h$ we have that:

$$\frac{1}{e^\epsilon} \leq \frac{M_{i,j}}{M_{h,j}} \leq e^\epsilon.$$

The a posteriori entropy of a channel with matrix M will be denoted by $H^M_\infty(A|B)$, and its min-entropy leakage by $I^M_\infty(A; B)$.

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

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

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

Also, given a partial function $\rho : \mathcal{A} \rightarrow \mathcal{B}$, 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".

Finally, given a graph $G = (\mathcal{V}, \sim)$ with diameter δ, we denote by Δ_G the set $\{0, 1, \ldots, \delta\}$. We may omit the subscript and denote the set only by Δ if the context does not allow for any confusion. For a fixed d, we define $n_d = |\mathcal{V}_{\langle d \rangle}(v)|$ as the number of vertices in $\mathcal{V}$ at distance d from v, and we intend that it will be always clear by the context to which set of vertices $\mathcal{V}$ and element v the value n_d is associated with.

4.2 The matrix transformation

Consider a channel whose matrix M has at least as many columns as rows and assume that the input distribution is uniform. Our transformation on the channel matrices is divided into two steps. First, we transform M into a matrix M' in which each of the first n columns has a maximum in the diagonal, and the remaining columns are all 0's. Second, under the assumption that the input domain is distance-regular or vertex-transitive, we transform M' into a matrix M'' whose diagonal elements are all the same, and coincide with the maximum element $\max^{M''}$ of M''. Our transformation ensures that both M' and M'' are valid channel matrices (i.e., each row is a probability distribution), respect ϵ-differential privacy, and preserve the value of the a posteriori entropy for the uniform input distribution. A scheme of our transformation is shown in Figure 5, where Lemma 1 (*Step* 1) is the first step of our transformation, and the second is step either Lemma 2 (*Step* 2*a*) or Lemma 4 (*Step* 2*b*), depending on whether the graph structure is distance-regular or vertex-transitive, respectively.

The next Lemma is relative to the first step, and it is independent of the graph structure of the input.

Lemma 1 (Step 1)**.** *Let M be a channel matrix of dimensions $n \times m$ s.t. $n \leq m$, and assume that M satisfies ϵ-differential privacy. Then it is possible to transform M into a matrix M' satisfying the following conditions:*

(i) *M' is a valid channel matrix: $\sum_{j=0}^{m-1} M'_{i,j} = 1$ for all $0 \leq i \leq n-1$;*

(ii) *Each of the first n columns has a maximum in the diagonal: $M'_{i,i} = \max_i^{M'}$ for all $0 \leq i \leq n-1$;*

(iii) *The $m-n$ last columns only contain 0's: $M'_{i,j} = 0$ for all $0 \leq i \leq n-1$ and all $n \leq j \leq m-1$;*

(iv) *M' satisfies ϵ-differential privacy: $\frac{M'_{i,j}}{M'_{h,j}} \leq e^{\epsilon}$ for all $0 \leq i, h \leq n-1$ s.t. $i \sim h$ and all $0 \leq j \leq m-1$;*

(v) *$H_\infty^{M'}(A|B) = H_\infty^{M}(A|B)$, if A has the uniform distribution.*

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

- $N_{i,\rho(i)} = \max_{\rho(i)}^{N}$ for all $i \in \mathcal{A}$, and

$$\begin{tikzcd}[ampersand replacement=\&]
M \begin{bmatrix} M_{0,0} & M_{0,1} & \dots & M_{0,m-1} \\ M_{1,0} & M_{1,1} & \dots & M_{1,m-1} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n-1,0} & M_{n-1,1} & \dots & M_{n-1,m-1} \end{bmatrix} \arrow[d, dashed, "\substack{\textit{Lemma}\text{ Step 1} \\ \text{(any graph structure)}}"] \\
M' \left[\begin{array}{cccc|ccc} \max_0^{M'} & - & \dots & - & 0 & \dots & 0 \\ - & \max_1^{M'} & \dots & - & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ - & - & \dots & \max_{n-1}^{M'} & 0 & \dots & 0 \end{array}\right] \arrow[d, dashed, bend right, "\substack{\textit{Lemma}\text{ Step }2a \\ \text{(distance-regular)}}"'] \arrow[d, dashed, bend left, "\substack{\textit{Lemma}\text{ Step }2b \\ \text{(vertex-transitive)}}"] \\
M'' \left[\begin{array}{cccc|ccc} \max^{M''} & - & \dots & - & 0 & \dots & 0 \\ - & \max^{M''} & \dots & - & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ - & - & \dots & \max^{M''} & 0 & \dots & 0 \end{array}\right]
\end{tikzcd}$$

Figure 5: Steps of the matrix transformation for distance-regular and vertex-transitive graphs

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

The first condition means that every row contains the maximum of exactly one column, and the second condition means that the exceeding columns are zero-ed columns.

We iteratively construct ρ and 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 $\rho_0(i) = \perp$ 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:

$$\begin{aligned} \rho_s &= \rho_{s-1} \cup \{i \mapsto j\} && \text{and} \\ N_s &= N_{s-1} \end{aligned}$$

2. $\rho_{s-1}(i) = k \in \mathcal{B}$. We "collapse" column j into column k:

$$\begin{aligned} \rho_s &= \rho_{s-1} && \text{and} \\ N_s &= N_{s-1}[j \to k] \end{aligned}$$

Since the operation of "collapsing" keeps the mapping $i \mapsto k$ in ρ_s and then zeroes the column j in N_s, all unassigned columns $\mathcal{B} \setminus \rho_m(\mathcal{A})$ must be zero in N_m. We finish the construction by taking ρ to be the same as ρ_m after assigning to every 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 $N_{i,\rho(i)} = \max^N_{\rho(i)}$ 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 (condition (iv)) because each column is a linear combination of columns of M. It is also easy to see that $\sum_j \max^N_j = \sum_j \max^M_j$, and from that it immediately follows that $H^N_\infty(A|B) = H^M_\infty(A|B)$ (recall that A has the uniform distribution and therefore the a posteriori entropy is a function of the sum of the maximum of each column), so condition (v) is satisfied.

Finally, we create our claimed matrix M' from N just by rearranging the columns according to ρ. Note that the order of the columns is irrelevant, since any permutation represents the same conditional probabilities and therefore the same channel[1]. The resulting matrix M' has all maxima in the diagonal $M'_{i,i}$ for $0 \leq i \leq n-1$, and every element in the columns $n \leq j \leq m-1$ are 0, which satisfies conditions (ii) and (iii). Also, since N is a valid channel matrix, so is M' and condition (i) is also satisfied.

□

The second step of our transformation depends on the graph structure of $(\mathcal{A}, \sim)$. Before we discuss this step, we need to introduce a notion of distance between elements in $\mathcal{B}$, derived from the notion of distance between elements in $\mathcal{A}$. Let M be a channel matrix in which the maximum of each column is in the diagonal, as in Figure 6. Then we define the distance between two elements $j_1, j_2 \in \mathcal{B}$ as follows:

$$d(j_1, j_2) = \begin{cases} d(i_1, i_2) & \text{if there exist } i_1, i_2 \in \mathcal{A} \text{ such that } i_1 = j_1 \text{ and } i_2 = j_2, \\ \bot & \text{otherwise.} \end{cases} \tag{3}$$

Note that the range of the notion of distance defined above is the set $\Delta = \{0, 1, \ldots, \delta\}$, where δ is the diameter of $(\mathcal{A}, \sim)$. Let $\mathcal{B}_{\langle d \rangle}(j) \subseteq \mathcal{B}$ be the set of elements at distance d from an element $j \in \mathcal{B}$. Of course, for every $j \in \mathcal{B}$, we have $\bigcup_{d \in \Delta} \mathcal{B}_{\langle d \rangle}(j) = \mathcal{B}$.

[1] By rearranging the columns of the channel matrix we may change the marginal probability of the outputs. This, however, poses no problem for our purposes, since the maximum a posteriori entropy of the channel will be maintained. If we want the marginal probability of the outputs to remain unchanged, we can just "re-label" the columns after the rearrangement so they will match the correct outputs.

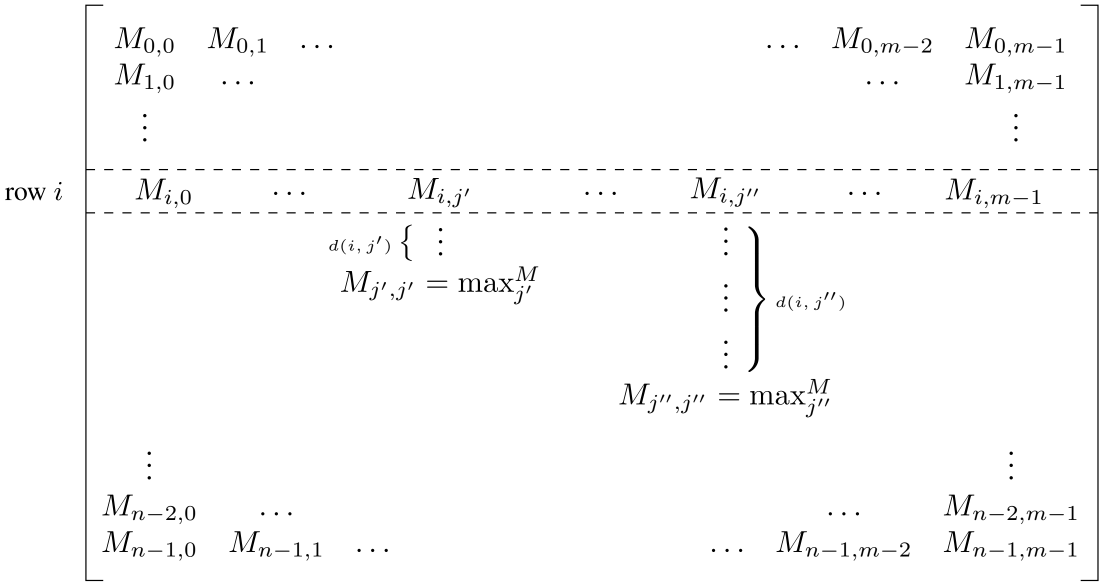


Figure 6: The relation between elements of a row i and the elements in the diagonal

Finally, let us extend the adjacency relation $\sim$ on $\mathcal{A}$ to an adjacency relation $\sim'$ on $\mathcal{B}$ by using the notion of distance of (3). For any $j_1, j_2 \in \mathcal{B}$, we have $j_1 \sim' j_2$ if and only if $d(j_1, j_2) = 1$. Let $\mathcal{B}^* = \{0, 1, \ldots, n-1\}$ be the subset of $\mathcal{B}$ that excludes the zero-ed columns of M' from n to $m-1$. Therefore, if $(\mathcal{A}, \sim)$ is distance-regular, so it is $(\mathcal{B}^*, \sim')$.

Now we are ready to present the lemma for the second step of our transformation, in the case of distance-regular graphs.

Lemma 2 (Step 2a)**.** *Let M' be a channel matrix of dimensions $n \times m$ such that $n \leq m$, and assume that M' satisfies ϵ-differential privacy. Let $\sim$ be an adjacency relation on $\mathcal{A}$ such that the graph $(\mathcal{A}, \sim)$ is connected and distance-regular. Assume that the maximum value of each column is on the diagonal, that is $M_{i,i} = \max_i^M$ for all $i \in \mathcal{A}$, and that all the last $m-n$ columns have only zero elements, i.e., $M'_{i,j} = 0$ for all $0 \leq i \leq n-1$ and $n \leq j \leq m-1$. Then it is possible to transform M' into a matrix M'' satisfying the following conditions:*

(i) *M'' is a valid channel matrix: $\sum_{j=0}^{m-1} M''_{i,j} = 1$ for all $0 \leq i \leq n-1$;*

(ii) *The elements of the diagonal are all the same, and are equal to the maximum of the matrix: $M''_{i,i} = \max^{M''}$ for all $0 \leq i \leq n-1$;*

(iii) *The $m-n$ last columns contain only 0's: $M''_{i,j} = 0$ for all $0 \leq i \leq n-1$ and all $n \leq j \leq m-1$;*

(iv) *M'' satisfies ϵ-differential privacy: $\frac{M'_{i,j}}{M'_{h,j}} \leq e^{\epsilon}$ for all $0 \leq i, h \leq n-1$ s.t. $i \sim h$ and all $0 \leq j \leq m-1$;*

(v) *$H_\infty^{M''}(A|B) = H_\infty^{M'}(A|B)$, if A has the uniform distribution.*

Proof. Let us consider $\mathcal{B}^* = \{0, 1, \ldots, n-1\}$, i.e., the subset of $\mathcal{B}$ that excludes the zero-ed columns of M' from n to $m-1$. Note that we can safely use the set $\mathcal{B}^*$ instead of $\mathcal{B}$ in this proof because the zero-ed columns do not contribute to the a posteriori entropy, and trivially respect ϵ-differential privacy.

We then define the matrix M'' as follows.

$$M''_{i,j} = \begin{cases} \frac{1}{n|\mathcal{A}_{\langle d(i,j)\rangle}(i)|} \sum_{k\in\mathcal{B}^*} \sum_{h\in\mathcal{A}_{\langle d(i,j)\rangle}(k)} M'_{h,k} & \text{if } j \in \mathcal{B}^*, \\ 0 & \text{otherwise.} \end{cases}$$

Condition (iii) is clearly satisfied, and we now show that M'' is indeed a channel matrix (condition (i))..

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

Note that for every i, $\mathcal{B}^* = \bigcup_{d\in\Delta} \mathcal{B}^*_{\langle d\rangle}(i)$, and for different values of d the sets $\mathcal{B}^*_{\langle d\rangle}(i)$ are disjoint. Therefore the summation over $j \in \mathcal{B}^*$ can be split as follows

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

since $\displaystyle\sum_{j\in\mathcal{B}^*_{\langle d\rangle}(i)} \frac{1}{|\mathcal{A}_{\langle d\rangle}(i)|} = 1$ (recall that the definition of distance in $\mathcal{B}$ is given as that in $\mathcal{A}$), we obtain

$$= \frac{1}{n} \sum_{k\in\mathcal{B}^*} \sum_{d\in\Delta} \sum_{h\in\mathcal{A}_{\langle d\rangle}(k)} M'_{h,k}$$

and now the summations over h and d can be joined together

$$= \frac{1}{n} \sum_{k\in\mathcal{B}^*} \sum_{h\in\mathcal{A}} M'_{h,k}$$

now reorganizing the summations, considering that M' is a channel matrix, and that $|\mathcal{B}^*| = n$ we have

$$\begin{aligned}
&= \frac{1}{n} \sum_{h \in \mathcal{A}} \sum_{k \in \mathcal{B}^*} M'_{h,k} \\
&= \frac{1}{n} \sum_{h \in \mathcal{A}} 1 \\
&= 1
\end{aligned}$$

which implies that condition (i) is satisfied.

We now turn our attention to the elements of the diagonal, which are all identical because:

$$M''_{i,i} = \frac{1}{n} \sum_{k \in \mathcal{B}^*} M'_{k,k}$$

To fulfill condition (ii) we still need to show that $M''_{i,i} = \max_i^{M''}$ for all $i \in \mathcal{A}$.

$$\begin{aligned}
M''_{i,j} &= \frac{1}{n|\mathcal{A}_{\langle d(i,j)\rangle}(i)|} \sum_{k \in \mathcal{B}^*} \sum_{h \in \mathcal{A}_{\langle d(i,j)\rangle}(k)} M'_{h,k} \\
&\leq \frac{1}{n|\mathcal{A}_{\langle d(i,j)\rangle}(i)|} \sum_{k \in \mathcal{B}^*} \sum_{h \in \mathcal{A}_{\langle d(i,j)\rangle}(k)} M'_{k,k} && \text{(since the biggest element is in the diagonal)} \\
&= \frac{1}{n} \sum_{k \in \mathcal{B}^*} M'_{k,k} \frac{1}{|\mathcal{A}_{\langle d(i,j)\rangle}(i)|} \sum_{h \in \mathcal{A}_{\langle d(i,j)\rangle}(k)} 1 \\
&= \frac{1}{n} \sum_{k \in \mathcal{B}^*} M'_{k,k} \frac{|\mathcal{A}_{\langle d(i,j)\rangle}(k)|}{|\mathcal{A}_{\langle d(i,j)\rangle}(i)|} \\
&= \frac{1}{n} \sum_{k \in \mathcal{B}^*} M'_{k,k} \cdot 1 && \text{(since the graph is distance-regular)} \\
&= M''_{i,i}
\end{aligned}$$

Since A has the uniform distribution, $H_\infty^{M'}(A|B) = H_\infty^{M''}(A|B)$ (condition (v)) follows immediately.

It remains to show that M'' satisfies ϵ-differential privacy (condition (iv)). We need to show that

$$M''_{i,j} \leq e^{\epsilon} M''_{i',j} \qquad \forall j \in \mathcal{B}, i, i' \in \mathcal{A} : i \sim i'$$

Note that for the columns $\mathcal{B} \backslash \mathcal{B}^*$ all elements are zero-ed, and therefore ϵ-differential privacy is trivially satisfied. Hence we only need to consider the elements in $\mathcal{B}^*$.

Since $d(i, i') = 1$, from the triangular inequality we have:

$$d(i', j) - 1 \leq d(i, j) \leq d(i', j) + 1$$

Thus, there are 3 possible cases:

1. $d(i, j) = d(i', j)$

   The result is immediate since $M''_{i,j} = M''_{i',j}$.

2. $d(i, j) = d(i', j) - 1$

   We define the set of neighbors of h “one step further away” from k:

   $$\mathcal{F}_{h,k} = \{h' \sim h \mid h' \in \mathcal{A}_{\langle d(h,k)+1\rangle}(k)\}$$

   Equivalently, we can see $\mathcal{F}_{h,k}$ as the set of neighbors of h that are not at distance $d(h, k) - 1$ from k. Since the graph is distance-regular, the number of elements in this set is given by the intersection number $b_{d(h,k)}$, i.e., $|\mathcal{F}_{h,k}| = b_{d(h,k)}$. The following inequalities hold for any $h, h' \in \mathcal{A}$:

   $$\begin{aligned} M'_{h,k} &\le e^{\epsilon} M'_{h',k} && \forall h' \in \mathcal{F}_{h,k} && \text{(diff. privacy)} \Rightarrow \\ b_{d(h,k)} M'_{h,k} &\le e^{\epsilon} \sum_{h' \in \mathcal{F}_{h,k}} M'_{h',k} && && \text{(sum of the above)} \end{aligned}$$

   fix now a distance d and sum the above inequalities for all vertices at distance d from h:

   $$\sum_{h \in \mathcal{A}_{\langle d\rangle}(k)} b_d M'_{h,k} \le e^{\epsilon} \sum_{h \in \mathcal{A}_{\langle d\rangle}(k)} \sum_{h' \in \mathcal{F}_{h,k}} M'_{h',k}$$

   Note that, being the graph distance-regular, and by the definition of the intersection number c_{d+1}, each $h' \in \mathcal{A}_{\langle d+1\rangle}(k)$ is contained in $\mathcal{F}_{h,k}$ for exactly c_{d+1} different $h \in \mathcal{A}_{\langle d\rangle}(k)$. So the right-hand side above sums all vertices of $\mathcal{A}_{\langle d+1\rangle}(k)$ exactly c_{d+1} times each. Thus we get that for all $k \in \mathcal{B}^*, d \in \Delta$:

   $$b_d \sum_{h \in \mathcal{A}_{\langle d\rangle}(k)} M'_{h,k} \le e^{\epsilon}\, c_{d+1} \sum_{h \in \mathcal{A}_{\langle d+1\rangle}(k)} M'_{h,k} \tag{4}$$

   Finally, note that $c_{d+1}|\mathcal{A}_{\langle d+1\rangle}(i)| = b_d|\mathcal{A}_{\langle d\rangle}(i)|$ (both sides count the number of edges between a vertex at distance d and a vertex at distance $d + 1$) and therefore we have

   $$\frac{c_{d+1}}{b_d} = \frac{|\mathcal{A}_{\langle d\rangle}(i)|}{\mathcal{A}_{\langle d+1\rangle}(i)|}$$

   And we conclude

   $$\begin{aligned} M''_{i,j} &= \frac{1}{n|\mathcal{A}_{\langle d\rangle}(i)|} \sum_{k \in \mathcal{B}^*} \sum_{h \in \mathcal{A}_{\langle d\rangle}(k)} M'_{h,k} \\ &\le e^{\epsilon} \frac{1}{n|\mathcal{A}_{\langle d\rangle}(i)|} \frac{c_{d+1}}{b_d} \sum_{k \in \mathcal{B}^*} \sum_{h \in \mathcal{A}_{\langle d+1\rangle}(k)} M'_{h,k} && \text{(from (4))} \\ &= e^{\epsilon} \frac{1}{n|\mathcal{A}_{\langle d+1\rangle}(i)|} \sum_{k \in \mathcal{B}^*} \sum_{h \in \mathcal{A}_{\langle d+1\rangle}(k)} M'_{h,k} \\ &= e^{\epsilon} M''_{i',j} \end{aligned}$$

3. $d(i, j) = d(i', j) + 1$

   This case is analogous to the case case where $d(i, j) = d(i', j) - 1$.

□

We will soon, in Lemma 4, cover the second step of our transformation for the case of vertex-transitive graphs. Before, however, we will derive an auxiliary result in Lemma 3, and for that we introduce the following notation.

Let $(V, \sim)$ be a graph, $v, w \in V$ be two of its vertices, and let Γ be an automorphism group for $(V, \sim)$. We define the set of automorphisms that map v to w as:

$$\Gamma_{v \mapsto w} = \{\sigma \in \Gamma \mid \sigma(v) = w\}$$

Note that $\Gamma_{v\mapsto w} \cap \Gamma_{v\mapsto w'} = \emptyset$ for all $w \neq w'$, and that $\Gamma = \bigcup_w \Gamma_{v\mapsto w}$ for any $v \in V$.

The following auxiliary lemma states that in a vertex-transitive graph, starting from any vertex v, there exist as many automorphisms mapping v to w as to any other vertex w'.

Lemma 3. *Let $(V, \sim)$ be a finite vertex transitive graph, $n = |V|$ and Γ its full automorphism group. Then, for all $v, w, w' \in V$:*

$$|\Gamma_{v\mapsto w}| = |\Gamma_{v\mapsto w'}| = \frac{|\Gamma|}{n}$$

Proof. Given two automorphisms σ and ρ, let $\rho \circ \sigma$ denote the composition of ρ with σ, i.e., the automorphism one would obtain by first applying σ to every vertex of the graph, and then applying ρ to the resulting mapping. By extension, given an automorphism group Γ and an automorphism ρ, we write $\rho \circ \Gamma$ for the automorphism group obtained by composing ρ with every $\sigma \in \Gamma$.

Assume that $|\Gamma_{v\mapsto w}| < |\Gamma_{v\mapsto w'}|$ for some $v, w, w' \in V$. Since the graph is vertex-transitive, there exists an automorphism ρ such that $\rho(w') = w$.

Consider the set of automorphisms $\rho \circ \Gamma_{v\mapsto w'}$. This set contains $|\Gamma_{v\mapsto w'}|$ distinct automorphisms (since $\rho \circ \sigma = \rho \circ \sigma'$ implies $\sigma = \sigma'$). Moreover these automorphisms map v to w, and therefore we have $\rho \circ \Gamma_{v\mapsto w'} \subseteq \Gamma_{v\mapsto w}$, which is a contradiction since by hypothesis we have $|\Gamma_{v\mapsto w}| < |\Gamma_{v\mapsto w'}| = |\rho \circ \Gamma_{v\mapsto w'}|$.

Thus $|\Gamma_{v\mapsto w}| \geq |\Gamma_{v\mapsto w'}|$ and by exchanging w, w' we get $|\Gamma_{v\mapsto w}| = |\Gamma_{v\mapsto w'}|$. Finally from $\Gamma = \bigcup_w \Gamma_{v\mapsto w}$ we get $|\Gamma_{v\mapsto w}| = \frac{|\Gamma|}{n}$.

□

Now we are ready to present the second step of our transformation for vertex-transitive graphs.

Lemma 4 (Step 2*b*)**.** *Consider a channel matrix M' satisfying the assumptions of Lemma 2, except for the assumption about distance-regularity: let instead $(\mathcal{A}, \sim)$ be vertex-transitive. Assume that the input and output domains of the channel are the same, i.e., $\mathcal{A} = \mathcal{B}$. Then it is possible to transform M' into a matrix M'' with the same properties as in Lemma 2.*

Proof. Let us consider $\mathcal{B}^* = \{0, 1, \ldots, n-1\}$, i.e., the subset of $\mathcal{B}$ that excludes the zero-ed columns of M' from n to $m-1$. Note that we can safely use the set $\mathcal{B}^*$ instead of $\mathcal{B}$ in this proof because the zero-ed columns do not contribute to the a posteriori entropy, and trivially respect ϵ-differential privacy.

Let Γ be the automorphism group of $(\mathcal{A}, \sim)$. Since $\mathcal{A} = \mathcal{B}$, it follows that Γ is also an automorphism group of $(\mathcal{B}, \sim')$ (the graph induced on outputs by the adjacency relation on inputs), and for all $i, j \in \mathcal{A}$ we define the elements of M'' as:

$$M''_{i,j} = \begin{cases} \frac{1}{|\Gamma|} \sum_{\sigma\in\Gamma} M'_{\sigma(i),\sigma(j)} & \text{if } j \in \mathcal{B}^*, \\ 0 & \text{otherwise.} \end{cases}$$

Condition (iii) is clearly satisfied, and we now show that M'' is indeed a channel matrix (condition (i)).

$$\begin{aligned} \sum_{j=0}^{n-1} M''_{i,j} &= \sum_{j=0}^{n-1} \frac{1}{|\Gamma|} \sum_{\sigma\in\Gamma} M'_{\sigma(i),\sigma(j)} \\ &= \frac{1}{|\Gamma|} \sum_{\sigma\in\Gamma} \sum_{j=0}^{n-1} M'_{\sigma(i),\sigma(j)} \\ &= \frac{1}{|\Gamma|} \sum_{\sigma\in\Gamma} 1 && \text{(since } M_{\sigma(i),\cdot} \text{ is a prob. distribution)} \\ &= 1 \end{aligned}$$

Then we show that $M''_{i,i} = M''_{j,j}$ for all $i, j \in \mathcal{A}$.

$$\begin{aligned} M''_{i,i} &= \frac{1}{|\Gamma|} \sum_{\sigma\in\Gamma} M'_{\sigma(i),\sigma(i)} \\ &= \frac{1}{|\Gamma|} \sum_{k=0}^{n-1} \sum_{\sigma\in\Gamma_{i\mapsto k}} M'_{\sigma(i),\sigma(i)} && \text{(since } \Gamma = \bigcup_k \Gamma_{i\mapsto k}\text{)} \\ &= \frac{1}{|\Gamma|} \sum_{k=0}^{n-1} \sum_{\sigma\in\Gamma_{i\mapsto k}} M'_{k,k} && (\sigma(i) = k \text{ since } \sigma \in \Gamma_{i\mapsto k}) \\ &= \frac{1}{|\Gamma|} \sum_{k=0}^{n-1} M'_{k,k} |\Gamma_{i\mapsto k}| \\ &= \frac{1}{|\Gamma|} \sum_{k=0}^{n-1} M'_{k,k} \frac{|\Gamma|}{n} && \text{(by Lemma 3)} \\ &= \frac{1}{n} \sum_{k=0}^{n-1} M'_{k,k} \end{aligned}$$

And we can conclude that every element in the diagonal of M'' is the same, as they are the average of the diagonal elements of M'. To fulfil condition (ii) we still need to show that $M''_{j,j} = \max_j^{M''}$ for all $j \in \mathcal{A}$.

$$\begin{aligned} M''_{j,j} &= \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} M'_{\sigma(j),\sigma(j)} \\ &\geq \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} M'_{\sigma(h),\sigma(j)} && \text{(since } M' \text{ respects } \epsilon\text{-d.p.)} \\ &= M'_{h,j} \end{aligned}$$

Then we show that M'' provides ϵ-differential privacy (condition (iv)). Let $i, h \in \mathcal{A}$ such that $i \sim h$. Note that $\sigma(i) \sim \sigma(h)$ for all $\sigma \in \Gamma$ since σ is an automorphism. Thus for all $j \in \mathcal{A}$ we have:

$$\begin{aligned} M''_{i,j} &= \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} M'_{\sigma(i),\sigma(j)} \\ &\leq \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} e^{\epsilon} M'_{\sigma(h),\sigma(j)} && (\epsilon\text{-diff. priv. of } M, \sigma(i) \sim \sigma(h)) \\ &= e^{\epsilon} M''_{h,j} \end{aligned}$$

Finally, we show that $H_\infty^{M''}(A) = H_\infty^{M'}(A)$ (condition (v)).

$$\begin{aligned} H_\infty^{M''}(A) &= \sum_{i=0}^{n-1} M''_{i,i} && \text{(the maxima of } M' \text{ are in the diagonal)} \\ &= \sum_{i=0}^{n-1} \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} M'_{\sigma(i),\sigma(i)} \\ &= \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} \sum_{i=0}^{n-1} M'_{\sigma(i),\sigma(i)} \\ &= \frac{1}{|\Gamma|} \sum_{\sigma \in \Gamma} H_\infty^{M'}(A) && \text{(the maxima of } M \text{ are also in the diagonal)} \\ &= H_\infty^{M'}(A) \end{aligned}$$

□

Note that in the proof above, the vertex-transitivity of the graph was only used to show that $M''_{i,i} = M''_{h,h}$. All the other points would hold for any graph, and using any set Γ of automorphisms to construct M''.

4.3 The bound on the a posteriori entropy of the channel

Once our transformation presented in the previous section has been applied, and the channel matrix respects the properties of M'', we use the graph structure of $(\mathcal{A}, \sim)$ to determine a bound on the a posteriori entropy $H_\infty^{M''}(A|B)$ of M''. Recall that our matrix transformation preserves the value of the a posteriori conditional entropy, so the bound we find is also valid for the original channel matrix we started with.

It is a known result in the literature (cfr. [8]) that, if the distribution of A is uniform, then the a posteriori min-entropy of the channel M is given by

$$H_\infty^M(A|B) = -\log_2 \frac{1}{n} \sum_{j \in \mathcal{B}} \max_j^M$$

Hence, under our assumption that the input distribution A is uniform, and knowing that the diagonal elements of the matrix M'' are all equal to the maximum $\max^{M''}$, we have

$$H_\infty^{M''}(A|B) = -\log_2 \max^{M''} \tag{5}$$

Therefore finding a bound on the a posteriori entropy of the channel M'' reduces to finding a bound on $\max^{M''}$. This is exactly what we do in this section.

We proceed by noting that the property of ϵ-differential privacy induces a relation between the ratio of elements at any distance:

Remark 1. Let M be a matrix satisfying ϵ-differential privacy. Then, for any column j, and any pair of rows i and h we have that:

$$\frac{1}{e^{\epsilon\, d(i,h)}} \leq \frac{M_{i,j}}{M_{h,j}} \leq e^{\epsilon\, d(i,h)}$$

In particular, since for every i it is the case that $M_{i,i} = \max^M$, then for each element $M_{i,j}$ we have that:

$$M_{i,j} \geq \frac{\max^M}{e^{\epsilon\, d(i,j)}} \tag{6}$$

which motivates the next proposition.

Proposition 1. *Let M be a channel matrix satisfying ϵ-differential privacy where for every $0 \leq i \leq n-1$ it is the case that $M_{i,i} = max^M$. Then:*

$$\max^M \leq \frac{1}{\sum_{d \in \Delta} \frac{n_d}{e^{\epsilon d}}}$$

Proof. The elements of a given row i of M represent a probability distribution, therefore they sum to 1.

$$\sum_j M_{i,j} = 1$$

By substituting (6) in the equation above we obtain:

$$\sum_j \left(\frac{\max^M}{e^{\epsilon d(i,j)}} \right) \leq 1$$

$$\sum_d \left(\frac{n_d}{e^{\epsilon d}} \max^M \right) \leq 1$$

and therefore

$$\max{}^{M} \leq \frac{1}{\sum_d \frac{n_d}{e^{\epsilon d}}}$$

□

Putting together all the steps of this section, we obtain our main result.

Theorem 1. *Consider a channel matrix M satisfying ϵ-differential privacy for some $\epsilon > 0$, assume that the inpjts distribution on A is uniform, and that $(\mathcal{A}, \sim)$ is either distance-regular or vertex-transitive. Then we have:*

$$H_\infty^M(A|B) \geq -\log_2 \frac{1}{\sum_d \frac{n_d}{e^{\epsilon\, d}}} \tag{7}$$

where $n_d = |\mathcal{A}_{\langle d \rangle}(i)|$ is the number of nodes $j \in \mathcal{A}$ at distance d from $i \in \mathcal{A}$.

Moreover, this bound it tight, in the sense that we can build a matrix for which (7) holds with equality.

Proof. The inequality follows directly from (5) and Proposition 1. To prove that the bound is tight, it is sufficient to define each element $M_{i,j}$ according to (6) with equality instead of inequality. □

In the next sections we will instantiate Theorem 1 to the channel from databases to reported answers, deriving a bound for the leakage of the mechanism (Section 5), and also to the sub-channel from true answers to reported answers, finding a bound for the utility (Sectoin 6).

5 Application to leakage

The correlation $\mathcal{L}(X, Z)$ between X and Z measures the information that the adversary can learn about the database by observing the reported answers. In this section we measure this information as min-entropy leakage, that is $\mathcal{L}(X, Z) = I_\infty(X; Z)$.

We shall consider that the input for the channel follows a uniform distribution, and derive bounds on information leakage imposed by differential privacy. Since the min-entropy leakage $I_\infty^M(X; Z)$ of a given matrix M is maximum when the input distribution is uniform [8], the bounds we present in this section (Theorem 2, Proposition 5, and Proposition 6) are valid for every possible input distribution. Moreover, since we model side information as input distributions, it follows that our bounds are valid for any possible side information the adversary may have.

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

As seen in Section 3, $\mathcal{K}$ can be modeled as a channel with input X and output Z. We will use here the bound on the posterior min-entropy of the channel obtained in Section 4. For that, we need to show first that the graph structure of the database domain $(\mathcal{X}, \sim)$ presents the symmetries we need.

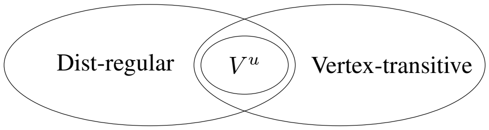


Figure 7: Venn diagram for the classes of graphs considered in Section 5.1.

5.1 Graph symmetries of the database domain

A *Hamming graph* is a graph where the vertices are tuples from some Cartesian space, and there is an edge between every two vertices if, and only if, they differ in the value of exactly one coordinate. Clearly the graph structure of the database domain $(\mathcal{X}, \sim)$ is Hamming graph, where each vertex is a database (i.e., a tuple containing the sensitive values of each individual), and there is an edge between every two databases that differ by the value of exactly one individual.

It is known in the literature [18, 9] that Hamming graphs present the two special kinds of symmetry we introduced in Section 2.3: distance-regularity and vertex-transitivity. The following propositions formalize this fact, which will be used in the remaining of this section to transform a generic channel matrix into a matrix with a symmetric structure, while preserving the a posteriori min-entropy and ϵ-differential privacy.

Proposition 2. *If $v \geq 2$, the graph $(\mathcal{V}^u, \sim)$ is a connected distance-regular graph with diameter $\delta = u$, and intersection numbers $b_d = (u - d)(v - 1)$ and $c_d = d$, for all $0 \leq d \leq \delta$.*

Proof. The vertices of $(\mathcal{V}^u, \sim)$ are u-tuples $(v_1, \ldots, v_u)$, $v_i \in \mathcal{V}$ and two vertices are adjacent if and only if the differ in exactly one element v_i. Then the distance between two vertices is the number of elements in which they differ. Let $x_1, x_2 \in \mathcal{V}^u$ with $d(x_1, x_2) = d$, so they differ in exactly d elements. To go at distance $d + 1$ from x_1 we can select any of the remaining $u - d$ elements and change it in $v - 1$ possible ways, so the total number is $(u - d)(v - 1)$, which only depends on d, not on x_1, x_2. Similarly, by changing one of the differing elements of x_2 to match the value of x_1 we get a vertex at distance $d - 1$, and there are d such elements. □

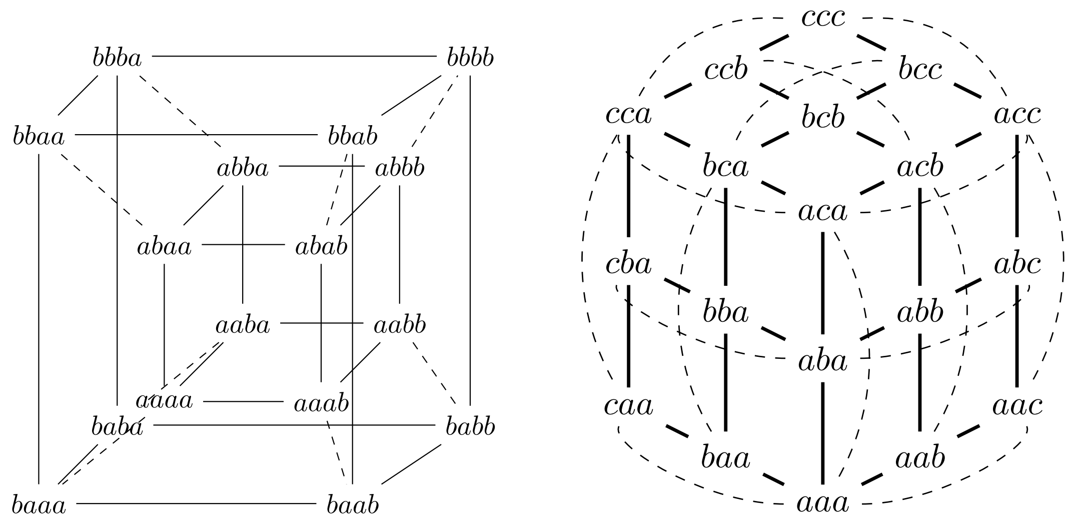
Proposition 3. *The graph $(\mathcal{V}^u, \sim)$ is a vertex-transitive graph.*

Proof. The Hamming graph is the Cartesian product of u complete graphs of size v. Complete graphs are trivially vertex-transitive and it is known [18] that a Cartesian product is vertex-transitive if, and only if, each of its factors is so. □

The relation between graph structures we consider in this paper is summarized in Figure 7. Figure 8 displays two database structures $(\mathcal{V}^u, \sim)$. Note that when $|\mathcal{V}| = 2$, $(\mathcal{V}^u, \sim)$ is the u-dimensional hypercube.

5.2 The bound on leakage

From Propositions 2 and 3 we know that $(\mathcal{X}, \sim)$ is both distance-regular and vertex-transitive, and therefore we can apply Theorem 1. Then, by (6) we know that for $j \in \mathcal{X}_{\langle d \rangle}(x)$ (i.e., every



(a) $u = 4, \mathcal{V} = \{a, b\}$ (4-dimensional hypercube)

(b) $u = 3, \mathcal{V} = \{a, b, c\}$ (for readability sake we only show part of the graph)

Figure 8: Two $(\mathcal{V}^u, \sim)$ graphs

j in $\mathcal{X}$ at distance d from a given x) it is the case that $M_{x,j} \geq \frac{\max^M}{e^{\epsilon d}}$. Furthermore we note that each element j at distance d from x can be obtained by changing the value of d individuals in the u-tuple representing i. We can choose these 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 x) in $v - 1$ possible ways. Therefore $|\mathcal{X}_{\langle d \rangle}(x)| = \binom{u}{d}(v-1)^d$, and we obtain that the number of databases at distance d from x is

$$n_d = |\mathcal{X}_{\langle d \rangle}(x)| = \binom{u}{d} (v-1)^d \tag{8}$$

Indeed, since x can be represented as a u-tuple with values in V, we need to select d individuals in the u-tuple and then change their values, and each of them can be changed in $v - 1$ different ways.

Using the value of n_d from (8) in Theorem 1 we obtain the following result.

Theorem 2. *If $\mathcal{K}$ satisfies ϵ-differential privacy, then the information leakage is bound from above as follows:*

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

Proof. For this proof we need a matrix with all column maxima on the diagonal, and all equal. We obtain such a matrix by transforming the matrix associated with $\mathcal{K}$ as follows: first we apply Lemma 1 to it (with $A = X$ and $B = Z$), and then we apply either Lemma 2 or Lemma 4 (we can choose either of them, since $(\mathcal{X}, \sim)$ is both distance-regular and vertex-transitive). The final matrix M has all non-zero elements on its $n \times n$ submatrix, with $n = |\mathcal{X}| = \mathcal{V}^u$, provides ϵ-differential privacy, and for every row i we have that $M_{i,i} = \max^M$. Furthermore, $I_\infty^M(X; Z)$ is equal to the min-entropy leakage of $\mathcal{K}$, assuming a uniform distribution on X.

Then we can derive:

$$
\begin{aligned}
\sum_{j=1}^{n} M_{i,j} &\geq \sum_{j=1}^{n} \frac{\max^M}{(e^\epsilon)^{d(i,j)}} && \text{(by (6))} \\
&= \sum_{d=0}^{u} n_d \frac{\max^M}{(e^\epsilon)^d} \\
&= \sum_{d=0}^{u} \binom{u}{d} (v-1)^d \frac{\max^M}{(e^\epsilon)^d} && \text{(by (8))}
\end{aligned}
$$

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

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

and by multiplying both sides of the inequality by $e^{\epsilon u}$ we get

$$\max^M \textstyle\sum_{d=0}^{u} \binom{u}{d} (v-1)^d e^{\epsilon(u-d)} \leq e^{\epsilon u}$$

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

$$\max^M \leq \left(\tfrac{e^\epsilon}{v-1+e^\epsilon}\right)^u \tag{9}$$

And hence, for the uniform input distribution:

$$
\begin{aligned}
I_\infty^M(X;Y) &= H_\infty(X) - H_\infty^M(X|Y) && \text{(by definition)} \\
&= \log_2 |\mathcal{V}^u| + \log_2 \max^M && \text{(by (5))} \\
&\leq \log_2 v^u + \log_2 \left(\frac{e^\epsilon}{v-1+e^\epsilon} \right)^u && \text{(by (9))} \\
&= u \log_2 \frac{v\, e^\epsilon}{v-1+e^\epsilon}
\end{aligned}
$$

To conclude our proof we recall that, the min-entropy leakage $I_\infty^M(X;Y)$ of a given matrix M is maximum when the input distribution is uniform [8]. Therefore since the above bound is valid for a uniform distribution on X, it is also valid for any distribution on X.

□

Note that the bound $Bnd(u, v, \epsilon) = u \log_2 \frac{v\, e^\epsilon}{(v-1+e^\epsilon)}$ is a continuous function in ϵ, has value 0 when $\epsilon = 0$, and converges to $u \log_2 v$ as ϵ approaches infinity. Figure 9 shows the growth of $Bnd(u, v, \epsilon)$ along with ϵ, for various fixed values of u and v.

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

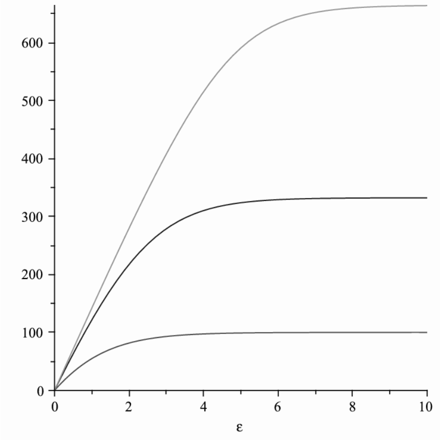


Figure 9: Graphs of $Bnd(u, v, \epsilon)$ for u=100 and v=2 (lowest line), v=10 (intermediate line), and v=100 (highest line), respectively.

Proposition 4. *For every u, v, and ϵ there exists a mechanism $\mathcal{K}$ which provides ϵ-differential privacy and whose min-entropy leakage, for the uniform input distribution, is $I_\infty(X; Z) = Bnd(u, v, \epsilon)$.*

Proof. The adjacency relation in $\mathcal{X}$ determines a graph structure $(\mathcal{X}, \sim)$. Let $\mathcal{Z} = \mathcal{X}$ and define the matrix of $\mathcal{K}$ as follows, where $d = d(x, z)$:

$$p(z|x) = \frac{2^{Bnd(u,v,\epsilon)}}{v^u (e^\epsilon)^d} \tag{10}$$

We need to show that $p(\cdot|x)$ is a probability distribution for every x:

$$\begin{aligned}
\sum_{z\in\mathcal{Z}} \frac{2^{Bnd(u,v,\epsilon)}}{v^u (e^\epsilon)^d} &= \frac{2^{Bnd(u,v,\epsilon)}}{v^u} \sum_{z\in\mathcal{Z}} \frac{1}{(e^\epsilon)^d} \\
&= \frac{2^{Bnd(u,v,\epsilon)}}{v^u} \sum_{d} \frac{n_d}{(e^\epsilon)^d} \\
&= \frac{2^{Bnd(u,v,\epsilon)}}{v^u} \cdot \frac{1}{\max^M} && \text{(by Proposition 1)} \\
&= \frac{2^{Bnd(u,v,\epsilon)}}{v^u} \cdot \frac{v^u e^0}{2^{Bnd(u,v,\epsilon)}} && \text{(max occurs when } d = 0 \text{ in (10))} \\
&= 1
\end{aligned}$$

We now show that $\mathcal{K}$ provides ϵ-differential privacy. For every $x, x' \in \mathcal{X}$ such that $x \sim x'$, and for every $z \in \mathcal{Z}$ we have:

$$\begin{aligned}
\frac{p(z|x)}{p(z|x')} &= \frac{2^{Bnd(u,v,\epsilon)}}{v^u (e^\epsilon)^{d(x,z)}} \cdot \frac{v^u (e^\epsilon)^{d(x',z)}}{2^{Bnd(u,v,\epsilon)}} && \text{(by (10))} \\
&= e^{\epsilon(d(x',z)-d(x,z))} \\
&\leq e^{\epsilon d(x,x')} && \text{(by the triangle inequality)} \\
&= e^\epsilon && (d(x, x') = 1 \text{ since } x \sim x')
\end{aligned}$$

Finally, let us calculate $I_\infty(X;Z) = H_\infty(X) - H_\infty(X|Z)$. Since input distribution is uniform, we have:

$$H_\infty(X) = -\log_2 \frac{1}{v^u}$$

Moreover, we know from (5) that $H_\infty(X|Z) = \max^{M''}$, and to obtain the maximum value we take $d = 0$ in (10):

$$\begin{aligned} H_\infty(X|Z) &= -\log_2 \frac{2^{Bnd(u,v,\epsilon)}}{v^u e^0} \\ &= -\log_2 \frac{2^{Bnd(u,v,\epsilon)}}{v^u} \end{aligned}$$

Now, by subtracting the value of $H_\infty(X|Z)$ from the value of $H_\infty(X)$ we obtain $I_\infty(X;Z) = Bnd(u,v,\epsilon)$.

□

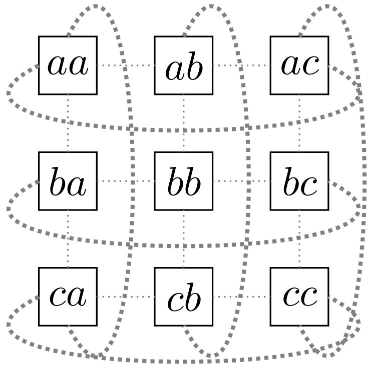
An example follows of the use of $Bnd(u,v,\epsilon)$ as a bound for the min-entropy leakage.

Example 1. Assume that we are interested in the eye color of a certain population $\mathcal{U} = \{Alice, Bob\}$. Let $\mathcal{V} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ where $\mathtt{a}$ stands for *absent* (i.e., the *null* value), $\mathtt{b}$ stands for *blue*, and $\mathtt{c}$ stands for *coalblack*. Each dataset is a tuple d_0d_1, where $d_0 \in \mathcal{V}$ represents the eye 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 10(a) represents the set $\mathcal{X}$ of all possible datasets and its adjacency relation. Fig 10(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 entry α stands for $\frac{\max^M}{e^{\epsilon\alpha}}$, where $\max^M$ is the highest value in the matrix, i.e., $\max^M = \frac{2^{Bnd(u,v,\epsilon)}}{v^u} = \left(\frac{ve^\epsilon}{v-1+e^\epsilon}\right)^u \frac{1}{v^u} = \frac{e^{2\epsilon}}{(2+e^\epsilon)^2}$.

Note that the bound $Bnd(u,v,\epsilon)$ is guaranteed to be reached with the uniform input distribution. The construction of the matrix for Proposition 4 gives a square matrix of dimension $\mathcal{V}^u \times \mathcal{V}^u$. 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 $r < \mathcal{V}^u$, and want to consider only those $\mathcal{K}$'s whose range has cardinality at most r. Proposition 5 shows that in this restricted setting we can find a better bound than the one given by Theorem 2. It relies on the following lemma.

Lemma 5. *Let $\mathcal{K}$ be a mechanism with input X, where $\mathcal{X} = \mathcal{V}^u$, providing ϵ-differential privacy. Assume that $r = |Range(\mathcal{K})| = v^\ell$, for some $\ell < u$. Let M be the matrix associated with $\mathcal{K}$. 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 $(\mathcal{V}^\ell, \sim_\ell)$, and such that:*

(i) M' is a valid channel matrix: $\sum_{j=0}^{m-1} M'_{i,j} = 1$ for all $0 \le i \le n-1$;



(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	2	1	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

Figure 10: All possible databases and highest min-entropy leakage matrix giving ϵ-differential privacy for Example 1.

(ii) $M'_{i,j} \leq (e^\epsilon)^{u-l+d}\, M'_{h,j}$ *for all* $i, h \in \mathcal{X}$ *and* $j \in \mathcal{Z}$*, where* d *is the* $\sim'$*-distance between* i *and* h*;*

(iii) The elements of the diagonal are all equal to the maximum element of the matrix: $M'_{i,i} = \max^{M'}$ *for all* $i \in \mathcal{X}$*;*

(iv) $H_\infty^{M'}(X|Z) = H_\infty^{M}(X|Z)$*, if* X *has the uniform distribution.*

Proof. (Sketch) 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 with any column.) Then define $\sim'$ as the projection of $\sim_u$ on $\mathcal{V}^\ell$, which satisfies condition (ii). Finally, apply the procedure in Lemma 2, or equivalently the procedure in Lemma 4, on the structure $(\mathcal{X}, \sim')$ to make all elements in the diagonal equal to the maximum element of the matrix (condition (iii)). Note that this procedure preserves the property of condition (ii), and conditional min-entropy ((iv)). Also the matrix obtained is a valid channel matrix (condition (i)). □

Now we are ready to prove the proposition.

Proposition 5. *Let* $\mathcal{K}$ *be a mechanism with associated channel matrix* M*, and let* $r = |Range(\mathcal{K})|$*. If* $\mathcal{K}$ *provides* ϵ-differential privacy *then the min-entropy leakage associated with* $\mathcal{K}$ *is bounded from above as follows:*

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

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

Proof. We first consider the case where $r = v^\ell$ for some ℓ. We transform the matrix M associated with $\mathcal{K}$ by applying Lemma 5, and let M' be the resulting matrix. Let $\max^{M'}$ be the value of every element in the diagonal of M', i.e., $\max^{M'} = M'_{i,i}$ for every row i, and let $\mathcal{A}'_{\langle d\rangle}(i)$ be the set of elements whose $\sim'$-distance from i is d. Note that for every $j \in \mathcal{A}'_{\langle d\rangle}(i)$ we have that $M'_{j,j} \leq M'_{i,j}(e^{\epsilon})^{u-\ell+d}$, hence

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

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 $(\mathcal{V}^{\ell}, \sim_{\ell})$). 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

$$|\mathcal{A}'_{\langle d\rangle}(i)| = \binom{\ell}{d}(v-1)^d$$

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

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

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

$$\max^{M} + \textstyle\sum_{d=1}^{\ell} \binom{\ell}{d}(v-1)^d \frac{\max^M}{(e^{\epsilon})^{u-\ell+d}} \;\leq\; 1 \tag{11}$$

Simple calculations, similar to those of the proof of Theorem 2, give:

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

Therefore:

$$\begin{aligned}
I_{\infty}^{M'}(X;Z) &= H_{\infty}(X) - H_{\infty}^{M'}(X|Z) && \text{(by definition)} && (12)\\
&= \log_2 v^u + \log_2 \sum_{j=1}^{v^{\ell}} \max^M \tfrac{1}{v^u} && && (13)\\
&= \log_2 v^u + \log_2 \frac{1}{v^u} + \log_2(v^{\ell} \max^M) && && (14)\\
&\leq \log_2 \frac{v^{\ell}\,(e^{\epsilon})^u}{(v-1+e^{\epsilon})^{\ell}-(e^{\epsilon})^{\ell}+(e^{\epsilon})^u} && \text{(by (11))} && (15)
\end{aligned}$$

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 $m = r - v^{\ell}$. Transform the matrix M associated with $\mathcal{K}$ by collapsing the m columns with the smallest maxima into the m columns with highest maxima. I.e., let $j_1, j_2, \ldots, j_m$ the indices of the columns which have smallest maxima values, i.e., $\max^M_{j_t} \leq \max^M_j$ 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 with maximum values. Recalling the definition of the "collapsing" operator introduced in Section 4.1, we 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

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

After transforming N into a matrix M' with the same min-entropy leakage as described in the first part of this proof, from (12) we conclude

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

□

Note that this bound can be much smaller than the one provided by Theorem 2. 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 $Bnd(u, v, \epsilon)$.

This does not contradict the fact that the bound $Bnd(u, v, \epsilon)$ is strict: in fact it is strict when we are free to choose the range, but here we fix the dimension of the range.

Leakage about an individual Protecting an entire database is not the primary 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.

The definition of differential privacy induces a straightforward bound on the information leakage about an individual. Let x be a database and $x^- = x \backslash \{d_i\}$ be the database obtained from x after removing the data relative to individual i. Of course $x \sim x^-$, and we can easily derive the following relation:

$$\begin{aligned}
\frac{p(z|x)}{p(z|x^-)} &= \frac{p(z)p(x|z)}{p(x)} \cdot \frac{p(x^-)}{p(z)p(x^-|z)} && \text{(by the Bayes law)}\\
&= \frac{p(x|z)}{p(x)} \cdot \frac{p(x^-)}{p(x^-|z)}\\
&= \frac{p(x^-|z)p(d_i|x^-, z)}{p(x^-)p(d_i|x^-)} \cdot \frac{p(x^-)}{p(x^-|z)} && \text{(by the chain rule of probabilities)}\\
&= \frac{p(d_i|x^-, z)}{p(d_i|x^-)} && (16)
\end{aligned}$$

Since by ϵ-differential privacy we have $\frac{p(z|x)}{p(z|x^-)} \leq e^\epsilon$, from (16) we conclude:

$$\frac{p(d_i|x^-, z)}{p(d_i|x^-)} \leq e^\epsilon \qquad (17)$$

In (17), $p(d_i|x^-)$ represents the probability of the adversary correctly guessing the value of d_i in one try if she knows the values of all other individuals in the database, x^-. Correspondingly, $p(d_i|x^-, z)$ represents the probability of a successful guess if the adversary knows x^- and has observed the output z of the differentially-private mechanism. The inequality shows that ratio between these two probabilities is no greater than e^ϵ, which is a natural bound on the min-entropy leakage of the mechanism (once we take logarithms).

With our model, however, we can find a slightly better bound. We start by fixing an individual, and we consider a tuple $x^- \in \mathcal{V}^{u-1}$ containing the given (and known) values of all other $u-1$ individuals. Then we create a channel whose input V ranges over the values in $\mathcal{V}$ and represents the value of our individual of interest. Note that this means that we take into consideration all possible input databases where the values of the other individuals are exactly those of x^- and only the value of the selected individual varies. Intuitively, $I_\infty^{x^-}(V;Z)$ measures the leakage about the individual's value in this case. (Similarly, $H_\infty^{x^-}(V|Z)$ represents the conditional entropy of V given Z for a fixed database where all other values are x^-.) Differential privacy provides a slightly stronger bound for leakage in this case.

In this context, the *leakage for a single individual* can be characterized as follows.

Proposition 6. *Assume that $\mathcal{K}$ satisfies ϵ-differential privacy. Then the information leakage for an individual is bound from above by:*

$$I_\infty^{x^-}(V;B) \leq \log_2 \frac{v\, e^\epsilon}{v-1+e^\epsilon}$$

Proof. Fix a database x, and a particular individual i in $\mathcal{U}$. The possible ways in which we can change the value of i in x are $v-1$. All the new databases obtained in this way are adjacent to each other, i.e., the graph structure associated with the input is a clique of v nodes. Recall that n_d is the number of elements of the input at distance d from a given element x. In this case we have

$$n_d = \begin{cases} 1 & \text{for } d = 0, \\ v-1 & \text{for } d = 1, \\ 0 & \text{otherwise.} \end{cases}$$

By substituting this value of n_d in Theorem 1, we get

$$\begin{aligned} H_\infty^{x^-}(V|Z) &\geq -\log_2 \frac{1}{1+\dfrac{v-1}{e^\epsilon}} \\ &= -\log_2 \frac{e^\epsilon}{v-1+e^\epsilon} \end{aligned} \tag{18}$$

The particular individual can present v different values, and thus in the case the input distribution is uniform its min-entropy is $H_\infty^{x^-}(V) = \log_2 v$.

$$\begin{aligned} I_\infty^{x^-}(V;Z) &= H_\infty^{x^-}(V) - H_\infty^{x^-}(V|Y) && \text{(by definition)} \\ &\leq \log_2 v + \log_2 \frac{e^\epsilon}{v-1+e^\epsilon} && \text{(by (18))} \\ &= \log_2 \frac{v\,e^\epsilon}{v-1+e^\epsilon} \end{aligned}$$

Since the min-entropy leakage is maximum in the case of the uniform input distribution, the result follows.

□

Note that the bound on the leakage for an individual does not depend on the size u of $\mathcal{U}$, nor on the database x^- that we fix. This is in accordance with the fact that the guarantees provided by differential-privacy are independent of the side information the adversary may have.

Finally, recall that Equation 17 is a natural bound of e^ϵ on the leakage about an individual imposed by differential privacy. Rewriting that bound in the notation used in this section, we have $I_\infty^{x^-}(V;B) \leq e^\epsilon$. Note that this differs from Proposition 6 up to a very slight factor of $\frac{v}{v-1+e^\epsilon}$.

6 Application to utility

In this Section we use the bounds we obtained on the a posteriori entropy of the channel to derive bounds on the utility of differentially-private mechanisms.

For our analysis we assume an oblivious mechanism. We can decompose the system into the cascade of two channels, and the utility becomes a property of the noise channel $\mathcal{H}$ mapping the true answer $y \in \mathcal{Y}$ into a reported answer $z \in \mathcal{Z}$ according to given probability distributions $\mathcal{H}_{y,z}$. The data analyst, however, does not necessarily take z as her guess for the true 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 data analyst can remap her guess to a value $y' \in \mathcal{Y}$ according to some strategy that maximizes her expected gain.

The standard way to define utility is by means of *gain* functions (see for instance [6]). We define $gain : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ and the value $gain(y, y')$ represents the reward for guessing the answer y' when the correct answer is y.

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

$$\mathcal{U}(Y,Z) = \sum_y p(y) \sum_{y'} p(y'|y) gain(y, y') \tag{19}$$

where $p(y)$ is the prior probability of true answer y, and $p(y'|y)$ is the probability of the data analyst guessing y' when the true answer is y. Naturally, the above equation has an implicit dependence on the reported answer z, as the data analyst will choose her guess y' based on the reported answer z she observes.

Assuming that the data analyst uses a remapping function $guess : \mathcal{Z} \to \mathcal{Y}$, we can derive the following characterization of the utility. Let $\delta_x(\cdot)$ represent the probability distribution which has value 1 on x and 0 elsewhere.

$$\begin{aligned}
\mathcal{U}(Y,Z) &= \sum_y p(y) \sum_{y'} p(y'|y) gain(y,y') && \text{(by (19))}\\
&= \sum_y p(y) \sum_{y'} \left(\sum_z p(z|y) p(y'|z) \right) gain(y,y') \\
&= \sum_y p(y) \sum_{y'} \left(\sum_z p(z|y) \delta_{y'}(guess(z)) \right) gain(y,y') && \text{(as } y' = guess(z))\\
&= \sum_y p(y) \sum_z p(z|y) \sum_{y'} \delta_{y'}(guess(z)) gain(y,y') \\
&= \sum_y p(y) \sum_z p(z|y) gain(y, guess(z)) \\
&= \sum_{y,z} p(y,z) gain(y, guess(z)) && (20)
\end{aligned}$$

The use of identity gain functions in the context of differential privacy was also investigated in [15][2]. We focus here on the so-called *identity* gain function, which is defined as

$$gain_{id}(y,y') = \begin{cases} 1 & \text{if } y = y', \\ 0 & \text{otherwise.} \end{cases} \qquad (21)$$

In the above equation, note that the value y' represents the data analyst's guess after the observed answer z, hence:

$$gain_{id}(y, guess(z)) = \delta_y(guess(z))$$

This kind of function represents the case in which there is no reason to prefer one answer over another, except if it is the *correct* answer. More precisely, we obtain some gain if and only if we guess the right answer. The answer domain can be equipped with a notion of distance (i.e., even if two answers are wrong, one of them may be "closer" to the correct one than the other), and the gain function could take into account the proximity of the reported answer to the true one. Naturally, in this case a "close" answer, even if wrong, is considered better than a "distant" one. We do not assume here a notion of distance, and therefore we will focus on the identity case.

By substituting *gain* with $gain_{id}$ in (20) we obtain:

$$\mathcal{U}(Y,Z) = \sum_{y,z} p(y,z) \delta_y(guess(z)) \qquad (22)$$

[2] The authors of [15] used the dual notion of *loss functions* instead of gain functions, but the final result is equivalent.

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

$$\begin{aligned}\mathcal{U}(Y, Z) &= \sum_z \max_y p(y, z) \\ &= \sum_z \max_y (p(y)\, p(z|y)) \qquad \text{(by the Bayes law)} \qquad (23)\end{aligned}$$

If the gain function is the identity, and the function *guess* is chosen to optimize utility (i.e., it represents the data analyst's best strategy), then there is a well-known correspondence between $\mathcal{U}$ and the Bayes risk / the a posteriori min-entropy. This correspondence is expressed by the following proposition:

Proposition 7. *Assume that function gain is the identity and the function guess is optimal. Then:*

$$\mathcal{U}(Y, Z) = \sum_z \max_y (p(y)\, p(z|y)) = 2^{-H_\infty(Y|Z)}$$

Proof. Just substitute (23) in the definition of conditional min-entropy: $H_\infty(Z \mid Y) = -\log_2 \sum_z \max_y((p(y)\, p(z|y))$. □

Note that while leakage is related to the ratio between the a priori and a posteriori min-entropy of the channel, utility concerns only the a posteriori min-entropy. This reflects the fact that leakage is a comparative measure, i.e., it tells how much the adversary's probability of success has improved with respect to an initial situation, whereas utility is an absolute measure, as it is only concerned about how much the mechanism's output tells about the true answer to the query.

6.1 The bound on the utility

Here we show that, in some special cases, the fact that $\mathcal{K}$ provides ϵ-differential privacy induces a bound on the utility defined in terms of a identity gain function. Recall that in Section 4.2 we extended the adjacency relation $\sim$ from the datasets $\mathcal{X}$ to the true answers $\mathcal{Y}$, in such a way that two values in $\mathcal{Y}$ are adjacent if they have pre-images that are adjacent. Intuitively, the function f associated with the query determines a partition on the set of all databases ($\mathcal{X} = \mathcal{V}^u$), and we say that two classes are adjacent if they contain an adjacent pair. For the sake of clarity, we will state formally the following direct consequence of (3):

Definition 4. Given $y, y' \in \mathcal{Y}$, with $y \neq y'$, we say that y and y' are adjacent (notation $y \sim y'$), if and only if there exist $x, x' \in \mathcal{V}^u$ with $x \sim x'$ such that $y = f(x)$ and $y' = f(x')$.

Since $\sim$ is symmetric on databases, it is also symmetric on $\mathcal{Y}$, therefore also $(\mathcal{Y}, \sim)$ forms an undirected graph.

Using the above concept of neighborhood for the inputs of the noise channel $\mathcal{H}$, we can show that in an oblivious mechanisms (see Figure 4) if the query f is deterministic, then the mechanism $\mathcal{K}$ provides ϵ-differential privacy with respect to neighbor databases if and only if $\mathcal{H}$ respects ϵ-differential privacy with respect to neighbor answers. Intuitively, this result follows from the fact that a deterministic query f remaps every database $x \in \mathcal{X}$ to a sole answer $y \in \mathcal{Y}$, working as a sort of "re-labeling" that substitutes databases for answers in the adjacency graph structure, and therefore preserving ϵ-differential privacy. Note also that if $\mathcal{K}$ is oblivious, the probability of any reported answer $z \in \mathcal{Z}$ does not depend on the database, but solely on the true answer y. Therefore under a deterministic f, two databases x and x' can be mapped to same value of y only if, for all z, $\mathcal{K}_{x,z} = \mathcal{K}_{x',z}$.

Proposition 8. *In an oblivious setting, if the query function f is deterministic, then the mechanism $\mathcal{K}$ satisfies ϵ-differential privacy with respect to every pair of neighbor databases $x, x' \in \mathcal{X}$ if and only if the noise channel $\mathcal{H}$ satisfies ϵ-differential privacy with respect to every pair of neighbor answers $y, y' \in \mathcal{Y}$.*

Proof. Since the matrix $\mathcal{K}$ can be obtained by the product of the two matrices corresponding to f and $\mathcal{H}$, we can derive that, for every pair of neighbor databases x and x' and for all reported answer z:

$$\begin{aligned}
\frac{p(z|x)}{p(z|x')} &= \frac{p(z|x)}{p(z|x')} \\
&= \frac{\sum_y p(y|x)p(z|y)}{\sum_y p(y|x')p(z|y)} && \text{(matrix multiplication)} \\
&= \frac{\sum_y \delta_{f(x)}(y)p(z|y)}{\sum_y \delta_{f(x')}(y)p(z|y)} && \text{(since } f \text{ is deterministic)} \\
&= \frac{p(z|f(x))}{p(z|f(x'))} && \text{(by def. of } \delta\text{)} \\
&= \frac{\mathcal{H}_{f(x),z}}{\mathcal{H}_{f(x'),z}}
\end{aligned}$$

It follows immediately that $\frac{\mathcal{K}_{x,z}}{\mathcal{K}_{x',z}} \leq e^{\epsilon}$ if and only if $\frac{\mathcal{H}_{f(x),z}}{\mathcal{H}_{f(x'),z}} \leq e^{\epsilon}$.

□

The link the above proposition establishes between the mechanism $\mathcal{K}$ and the noise channel $\mathcal{H}$ will help us determine a bound on the utility of $\mathcal{H}$, since, in the case the query f is deterministic, requiring $\mathcal{K}$ to respect ϵ-differential privacy is equivalent to requiring that $\mathcal{H}$ does.

Theorem 3. *Consider a noise channel $\mathcal{H}$ satisfying ϵ-differential privacy for some $\epsilon > 0$, and let y be an element of $\mathcal{Y}$. Assume that the distribution of Y is uniform and that $(\mathcal{Y}, \sim)$ is either* *distance-regular or vertex-transitive. For each distance $d \in \{0, 1, \ldots, \delta\}$, where δ is the diameter of $(\mathcal{Y}, \sim)$, we have:*

$$\mathcal{U}(Y, Z) \leq \frac{1}{\sum\limits_{d} \frac{n_d}{e^{\epsilon\, d}}} \tag{24}$$

where n_d is the number of nodes $y' \in \mathcal{Y}$ at distance d from y.

Proof. Since $(\mathcal{Y}, \sim)$ is distance-regular or vertex-transitive, we can apply Theorem 1 to derive that $H_\infty^M(Z|Y) \geq -\log_2 \frac{1}{\sum_d \frac{n_d}{e^{\epsilon\, d}}}$. Then we just substitute this result in Proposition 7. □

Provided $(\mathcal{Y}, \sim)$ is distance-regular or vertex-transitive, the above bound is tight. Indeed, we can construct a noise channel $\mathcal{H}$ which satisfies (24) with equality. More precisely, for $0 \leq i \leq n-1$ and $0 \leq j \leq n-1$, we define $\mathcal{H}$ (here identified with its channel matrix for simplicity) as follows:

$$\mathcal{H}_{i,j} = \frac{\gamma}{e^{\epsilon\, d(i,j)}} \tag{25}$$

where

$$\gamma = \frac{1}{\sum\limits_{d} \frac{n_d}{e^{\epsilon\, d}}} \tag{26}$$

Note that $\mathcal{H}$ is a square matrix of dimension $n \times n$, where $n = |\mathcal{X}|$. This is not a problem because since we assume $(\mathcal{Y}, \sim)$ to be either distance-regular or vertex-transitive, via Theorem 1 we can transform the channel matrix into an equivalent one such that all non zero elements are in the submatrix of dimensions $n \times n$. Let $\mathcal{Z}^* = \{0, 1, \ldots, n-1\}$ be the subset of $\mathcal{Z}$ that excludes the zero-ed columns of the channel matrix from n to $m-1$.

Theorem 4. *Assume $(\mathcal{Y}, \sim)$ is distance-regular or vertex-transitive and that the distribution of Y is uniform. Then the matrix $\mathcal{H}$ defined in (25) satisfies ϵ-differential privacy and has maximal utility:*

$$\mathcal{U}(Y, Z) = \frac{1}{\sum\limits_{d} \frac{n_d}{e^{\epsilon\, d}}}$$

Proof. First we prove that the matrix as defined in (25) is a channel matrix, i.e., that each row is a probability distribution.

$$\begin{aligned}
\sum_{j\in\mathcal{Z}^*}\mathcal{H}_{i,j} &= \sum_{j\in\mathcal{Z}^*}\frac{\gamma}{e^{\epsilon d(i,j)}} \\
&= \gamma\sum_{j\in\mathcal{Z}^*}\frac{1}{e^{\epsilon d(i,j)}} \\
&= \gamma\sum_{d}\frac{n_d}{e^{\epsilon d}} && \text{by (26)} \\
&= \gamma\frac{1}{\gamma} \\
&= 1
\end{aligned}$$

Now we show that the utility is maximum.

$$\begin{aligned}
\mathcal{U}(Y,Z) &= \sum_{z\in\mathcal{Z}^*}\max_y(p(y)\,\mathcal{H}(z|y)) && \text{by (23)} \\
&= \sum_{z\in\mathcal{Z}^*}\max_y\frac{1}{|\mathcal{Y}|}\mathcal{H}(z|y) && \text{since } Y \text{ is uniform} \\
&= \frac{1}{|\mathcal{Y}|}\sum_{z\in\mathcal{Z}^*}\max_y\frac{\gamma}{\max_d e^{\epsilon d(z,y)}} && \text{by (25)} \\
&= \frac{1}{|\mathcal{Y}|}\sum_{z\in\mathcal{Z}^*}\gamma && \text{maximum is } d = 0 \\
&= \frac{1}{|\mathcal{Y}|}\cdot|\mathcal{Z}^*|\gamma \\
&= \gamma && \text{since } |\mathcal{Y}| = |\mathcal{Z}^*| = n
\end{aligned}$$

□

It is possible to always have $\mathcal{H}$ as in (25): the matrix so defined will be a legal channel matrix, and it will satisfy ϵ-differential privacy. If $(\mathcal{Y},\sim)$ is neither distance-regular nor vertex-transitive, then the utility of such $\mathcal{H}$ is not necessarily optimal.

6.2 Examples

The conditions for the construction of the optimal noise channel are strong, but there are some interesting scenarios in which they are satisfied. Depending on the degree of connectivity c of the graph $(\mathcal{Y},\sim)$, 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 *clique*, i.e., every element has exactly $|\mathcal{Y}| - 1$ adjacent elements.

- $(\mathcal{Y}, \sim)$ is a *ring*, i.e., every element has exactly two adjacent elements. This is similar to the case of the counting queries considered in [15], with the difference that our "counting" is in arithmetic modulo $|\mathcal{Y}|$.

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 as to meet those conditions. For computing the distance in (25) we use, instead of $(\mathcal{Y}, \sim)$, a structure $(\mathcal{Y}, \sim')$ which satisfies the conditions of Theorem 4, and such that $\sim \subseteq \sim'$. The matrix constructed in this way provides ϵ-differential privacy, but in general is not optimal. It is clear that, in general, the smaller $\sim'$, the higher the utility.

The matrices generated by (25) 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 each entry corresponds to a voter and contains the following three fields:

- *Id*: a unique (anonymized) identifier assigned to each voter;
- *City*: the name of the city where the individual voted;
- *Candidate*: the name of the candidate the individual voted for.

Consider the query *"What is the city with the greatest number of votes for a given candidate cand?"*. For such a query the identity gain function could be taken as the natural choice: from the data analyst's point of view, only the right city could give some gain, and all wrong answers would be equally bad. Also, since every two answers are neighbors, the graph structure of the answers is a clique.

Consider the scenario where the set of cities is $City = \{A, B, C, D, E, F\}$ 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 candidate *cand*. Table 1 shows two alternative mechanisms providing ϵ-differential privacy (with $\epsilon = \ln 2$). The first one, M_1, is based on the truncated geometric mechanism method used in [15] for counting queries (here extended to the case where every two distinct answers are neighbors). The second mechanism, M_2, is obtained by applying the definition of (25). From Theorem 4 we know that for the uniform input distribution M_2 gives optimal utility.

For the uniform input distribution, we have $\mathcal{U}(M_1) = 0.2243 < 0.2857 = \mathcal{U}(M_2)$. 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 $\mathcal{U}(M_1) = 0.2415 < 0.2857 = \mathcal{U}(M_2)$. This is not too surprising: the geometric mechanism, as well as the Laplacian mechanism proposed by Dwork, perform very well when the domain of answers is provided with a metric and the utility function is not the identity.

(a) M_1: adapted truncated geometric mechanism

In/Out	*A*	*B*	*C*	*D*	*E*	*F*
A	0.534	0.060	0.053	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	0.053	0.060	0.534

(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 candidate *cand*

Example 3. Consider the same database of the previous example, but now assume a counting query of the form *"What is the number of votes for candidate cand?"*. Now each answer has at most two neighbors, and the graph structure on the answers is a line. For illustration purposes, 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 [15] and the mechanism we propose M_2. Note that in order to apply our method we have first to apply Remark 2 to transform the graph structure from a line into a ring.

Consider the uniform prior distribution. We see that the utility of M_1 is higher than the utility of M_2, the first being $4/9$ and the second being $8/21$. This does not contradict our theorem, because our matrix is guaranteed to be optimal only in the case of a ring structure, not of 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 [15].

7 Related work

To the best of our knowledge, the first work to investigate the relation between differential privacy and information-theoretic leakage *for an individual* was [3]. There, the definition of channel was for a given database, and the channel inputs were all possible databases adjacent to it. Two bounds on leakage were presented, one for the min-entropy, and one for Shannon entropy. Our bound in Proposition 6 improves the min-entropy bound of [3].

Barthe and Köpf [5] were the first to investigate the more challenging connection between differential privacy and the min-entropy leakage *for the set of all possible databases*. They

(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	8/21	4/21	2/21	1/21	2/21	4/21
1	4/21	8/21	4/21	2/21	1/21	2/21
2	2/21	4/21	8/21	4/21	2/21	1/21
3	1/21	2/21	4/21	8/21	4/21	2/21
8	2/21	1/21	2/21	4/21	8/21	4/21
5	4/21	2/21	1/21	2/21	4/21	8/21

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

considered "end-to-end differentially-private mechanisms", which correspond to what we call the mechanism $\mathcal{K}$ in this paper, and proposed, like we do, to interpret these mechanisms as information-theoretic channels. They gave a bound for the leakage, but pointed out that it was not tight in general. They also observed that for any number of individuals u and level of privacy ϵ one can construct a channel whose maximal leakage is $u \log(2e^{\epsilon})/(e^{\epsilon}+1)$ and concluded therefore that the bound must be at least as high as such expression. Another difference between their work and ours is that [5] captures the case in which the focus of differential privacy is on hiding *participation* of individuals in a database, whereas we consider both the participation and the *values* of the individuals.

Clarkson and Schneider also considered differential privacy as a case study of their proposal for quantification of integrity [10]. There, the authors analyzed database privacy conditions from the literature (such as differential privacy, k-anonymity, and l-diversity) for utility quantification. In particular, they studied the relationship between differential privacy and a notion of leakage (being different from ours as their definition is based on Shannon entropy) and they provided a tight bound on leakage.

Heusser and Malacaria [17] 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 statistical analysis of the information leaked by the query. However, [17] did not attempt to relate information leakage to differential privacy.

In [15] the authors aimed at obtaining optimal-utility 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 in the sense it provides a general definition of utility for a mechanism M that captures any possible side information and preference 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 6 do not restrict to counting queries, but on the other hand we only consider the case of identity loss function.

Finally, our definition of the channel matrix in (25) corresponds to the exponential mechanism of McSherry and Talwar [21] if the quality function relating two answers i and j is taken to be $-d(i, j)$, that is, the negative of the distance between them. Therefore, under the hypothesis of Theorem 3, it follows that the exponential mechanism is an optimal way to maximize utility as measured by the identity gain-function, while preserving differential privacy.

8 Paper summary and discussion

In this paper we have investigated the relations of ϵ-differential privacy with leakage, and utility, extending our previous work [1, 2]. Our main contribution has been the development of a general technique for determining these relations depending on the graph structure of the input domain, induced by the adjacency relation and by the query. We have considered two particular structures, the distance-regular graphs, and the vertex-transitive graphs, which allowed us to obtain tight bounds on leakage and on utility. We also constructed an optimal noise channel satisfying ϵ-differential privacy for some special cases.

As future work, we plan to extend our result to other kinds of utility functions, and we believe that the g-leakage framework [4] may be suitable for this goal. In particular, we are interested in the case in which the the answer domain is provided with a metric, and we are interested in taking into account the degree of accuracy of the inferred answer.

Bibliography

[1] Mário S. Alvim, Miguel E. Andrés, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catuscia Palamidessi. Differential Privacy: on the trade-off between Utility and Information Leakage. In Gilles Barthe, Anupam Datta, and Sandro Etalle, editors, *Postproceedings of the 8th International Worshop on Formal Aspects in Security and Trust (FAST)*, volume 7140 of *Lecture Notes in Computer Science*, pages 39–54, Leuven, Belgium, March 2011. Springer.

[2] Mário S. Alvim, Miguel E. Andrés, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. On the relation between Differential Privacy and Quantitative Information Flow. In Jiri Sgall Luca Aceto, Monika Henzinger, editor, *38th International Colloquium on Automata, Languages and Programming (ICALP)*, volume 6756 of *Lecture Notes in Computer Science*, pages 60–76, Zurich, Switzerland, 2011. Springer.

[3] Mário S. Alvim, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catuscia Palamidessi. Differential privacy versus quantitative information flow. Technical report, INRIA and LIX, Ecole Polytechnique, 2010. http://hal.archives-ouvertes.fr/hal-00548214/en/.

[4] Mário S. Alvim, Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Geoffrey Smith. Measuring information leakage using generalized gain functions. In *Proceedings of the 25th IEEE Computer Security Foundations Symposium (CSF)*, pages 265–279, 2012.

[5] Gilles Barthe and Boris Köpf. Information-theoretic bounds for differentially private mechanisms. In *Proceedings of the 24th IEEE Computer Security Foundations Symposium (CSF)*, pages 191–204. IEEE Computer Society, 2011.

[6] Jose M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. John Wiley & Sons, Inc., 1994.

[7] Avrim Blum, Katrina Ligett, and Aaron Roth. A learning theory approach to non-interactive database privacy. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC)*, pages 609–618. ACM, 2008.

[8] Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Quantitative notions of leakage for one-try attacks. In *Proceedings of the 25th Conf. on Mathematical Foundations of Programming Semantics*, volume 249 of *Electronic Notes in Theoretical Computer Science*, pages 75–91. Elsevier B.V., 2009.

[9] A.E. Brouwer, A.M. Cohen, and A. Neumaier. *Distance Regular Graphs*. Ergebnisse der Mathematik 3.18. Springer-Verlag, 1989.

[10] M. R. Clarkson and F. B. Schneider. Quantification of integrity. *Mathematical Structures in Computer Science*, 2011. To appear.

[11] Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *33rd International Colloquium on Automata, Languages and Programming (ICALP)*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.

[12] Cynthia Dwork. Differential privacy in new settings. In Moses Charikar, editor, *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 174–183. SIAM, 2010.

[13] Cynthia Dwork. A firm foundation for private data analysis. *Communications of the ACM*, 54(1):86–96, 2011.

[14] Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC)*, pages 371–380, Bethesda, MD, USA, May 31 - June 2 2009. ACM.

[15] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In *Proceedings of the 41st annual ACM Symposium on Theory of Computing (STOC)*, pages 351–360, New York, NY, USA, 2009. ACM.

[16] Moritz Hardt and Guy N. Rothblum. A multiplicative weights mechanism for privacy-preserving data analysis. In *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, pages 61–70, Washington, DC, USA, 2010. IEEE Computer Society.

[17] Jonathan Heusser and Pasquale Malacaria. Applied quantitative information flow and statistical databases. In Pierpaolo Degano and Joshua D. Guttman, editors, *Proceedings of the International Workshop on Formal Aspects in Security and Trust (FAST 2009)*, volume 5983 of *Lecture Notes in Computer Science*, pages 96–110. Springer, 2009.

[18] Wilfried Imrich and Sandi Klavžar. *Product graphs, structure and recognition*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, 2000.

[19] Shiva Prasad Kasiviswanathan and Adam Smith. A note on differential privacy: Defining resistance to arbitrary side information. Cryptology ePrint Archive, Report 2008/144, 2008. http://eprint.iacr.org/.

[20] Boris Köpf and David A. Basin. An information-theoretic model for adaptive side-channel attacks. In Peng Ning, Sabrina De Capitani di Vimercati, and Paul F. Syverson, editors, *Proceedings of the 2007 ACM Conference on Computer and Communications Security (CCS 2007)*, pages 286–296. ACM, 2007.

[21] Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007)*, pages 94–103. IEEE Computer Society, 2007.

[22] Aaron Roth and Tim Roughgarden. Interactive privacy via the median mechanism. In *Proc. of the 42nd ACM Symposium on Theory of Computing (STOC)*, pages 765–774, 2010.

[23] Geoffrey Smith. On the foundations of quantitative information flow. In Luca de Alfaro, editor, *Proceedings of the 12th International Conference on Foundations of Software Science and Computation Structures (FOSSACS 2009)*, volume 5504 of *LNCS*, pages 288–302, York, UK, 2009. Springer.

[24] Geoffrey Smith. Quantifying information flow using min-entropy. In *Eighth International Conference on Quantitative Evaluation of Systems, (QEST), Aachen, Germany, 5-8 September, 2011*, pages 159–167, 2011.

MEALS Partner Abbreviations

SAU: Saarland University, D

RWT: RWTH Aachen University, D

TUD: Technische Universität Dresden, D

INR: Institut National de Recherche en Informatique et en Automatique, FR

IMP: Imperial College of Science, Technology and Medicine, UK

ULEIC: University of Leicester, UK

TUE: Technische Universiteit Eindhoven, NL

UNC: Universidad Nacional de Córdoba, AR

UBA: Universidad de Buenos Aires, AR

UNR: Universidad Nacional de Río Cuarto, AR

ITBA: Instituto Técnológico Buenos Aires, AR