# Quantitative Information Flow and Applications to Differential Privacy

Mário Alvim, Miguel Andrés, Konstantinos Chatzikokolakis, Catuscia Palamidessi

# Quantitative Information Flow and applications to Differential Privacy*

Mário S. Alvim, Miguel E. Andrés,
Konstantinos Chatzikokolakis, and Catuscia Palamidessi

INRIA and LIX, Ecole Polytechnique, France.

**Abstract.** Secure information flow is the problem of ensuring that the information made publicly available by a computational system does not leak information that should be kept secret. Since it is practically impossible to avoid leakage entirely, in recent years there has been a growing interest in considering the quantitative aspects of information flow, in order to measure and compare the amount of leakage. Information theory is widely regarded as a natural framework to provide firm foundations to quantitive information flow. In this notes we review the two main information-theoretic approaches that have been investigated: the one based on Shannon entropy, and the one based on Rényi min-entropy. Furthermore, we discuss some applications in the area of privacy. In particular, we consider statistical databases and the recently-proposed notion of differential privacy. Using the information-theoretic view, we discuss the bound that differential privacy induces on leakage, and the trade-off between utility and privacy.

## 1 Introduction

In the last few decades the amount of information flowing through computational systems has increased dramatically. Never before in history has a society been so dependent on such a huge amount of information being generated, transmitted and processed. It is expected that this vertiginous trend of increase will continue in the near future, reinforcing the need for efficient and safe ways to cope with this reality.

One of the concerns in the use of computational systems is 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 in practice this goal is usually impossible to achieve. The following example illustrates some of the issues.

* This work has been partially supported by the project ANR-09-BLAN-0169-01 PANDA and by the INRIA DRI Equipe Associée PRINTEMPS. The work of Miguel E. Andrés has been supported by the LIX-Qualcomm postdoc fellowship.

*Example 1.* Consider the password-checker algorithm defined in Table 1, where $K_1K_2 \dots K_N$ is the sequence of $N$ digits that compose a password, and $x_1x_2 \dots x_N$ is the string entered by the user. An attacker may obtain the password by simply trying a string, if she is so lucky to guess the string that matches the password.

```
out := OK
for i = 1, ..., N do
  if x_i ≠ K_i then
    out := FAIL
    exit()
  end if
end for
```

**Table 1.** Password-checker pseudocode.

Furthermore, even if the attacker makes the wrong guess, she still obtains some (small) amount of information: the information that the password is not the string she just entered, thus restricting the guessing range of potential further attempts.

Worse yet, the algorithm is subject to timing attacks: the attacker may be able to determine from the duration of the execution how many iterations are performed, thus inferring a prefix of the password even in case of failure.

Thus it is important to express the amount of leakage in quantitative terms, so to be able to assess whether a system is better than another, although they may both be insecure.

Another reason to consider the quantitative aspects is that the system may have a probabilistic behavior. This is the case, for instance, of *anonymity protocols*, which use randomization to obfuscate the link between the identity of the *culprit* and the observable outcome.

*Example 2.* DC-Net is a protocol for anonymous broadcasting based on the paradigm of the dining cryptographers [8].

The participants in the protocol are assumed to be the vertexes of a graph, whose edges are associated to binary coins which are visible to the adjacent vertexes. The protocol works as follows: suppose that a participant $x$ wants to broadcast one bit $b$. Then, all coins get flipped, and each participant reads the value of the coins which are visible to her and computes the binary sum of these values. Then all participants but $x$ declare the result, while $x$ (binarily) adds $b$ to her result, and then declares it. Fig. 1 illustrates the situation, where the labels of the edges represent the outcome of the coin tosses, and the labels of the vertices represent the declaration of each user (i.e the binary sum they calculate).

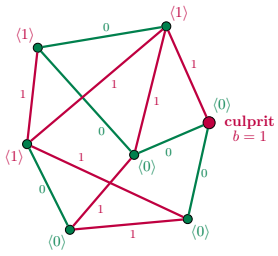



**Fig. 1.** Example of a dc-net.

It is easy to see that, since each coin is counted twice, the global contribution of all coins is 0. Hence the binary sum of all the declarations gives $b$, thus achieving the goal of making $b$ public. Furthermore thanks to the "noise" created by the coins, it is generally impossible for an adversary to determine with certainty the culprit, i.e. the vertex who is broadcasting $b$. We will see that a stronger property actually holds: if the coins are fair, and the graph is connected, then

for an external adversary the probability of each vertex to be the culprit does not change after she observes the declarations.

Several authors have proposed to use concepts from information theory to model information flow and to define the leakage in a quantitative way. So far, most of the approaches were based on *Shannon entropy* [23, 22, 28, 9, 19, 20, 6]. This is probably due to the fact that Shannon entropy is the most established and useful notion of entropy in information theory, because of its mathematical properties and its correspondence with the channel transmission rate. Nevertheless, other notions based on information theory have been considered, and argued to be more appropriate for security in certain scenarios. These include: *Rényi min-entropy* [25, 27], *guessing entropy* [21], and *marginal guesswork* [24].

The common idea in these information-theoretic approaches 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 easy is for an attacher to discover the secret. Therefore, as argued in [**?**], the notion of entropy should be chosen according to the model of attacker, and to the way we extimate the success of the attack. Normally, the entropy of a random variable represents its *uncertainty*, hence the vulnerability is anti-monotonic on the entropy.

Independently from the intended model of attacker, 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:

$$\textit{information leakage} = \textit{initial uncertainty} - \textit{remaining uncertainty} \quad (1)$$

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 all the notions of entropy mentioned above, when the channel is deterministic. In the more general case of a probabilistic channel, however, the non-negativeness is ensured only for Shannon entropy and for Rényi min-entropy. In these notes we aim at encompassing also the probabilistic case, hence we will restrict to the latter two notions of entropy.

### 1.1 Application to privacy

Recently, the above notions have been applied also to the problem of *statistical disclosure control* in the area of databases. The goal is to learn properties of the population as a whole, while maintaining the privacy of individuals. For example, in medical research it is desirable to collect the personal medical information of a large number of individuals. Researchers or public authorities can calculate a series of statistics from the sample and decide, for instance, how much money the health care system should spend next year in the treatment of a specific

disease. It desirable, however, that the participation in the sample will not harm the privacy of any individual: usually people do not want to have disclosed their specific status with respect to a given disease, or other personal matters. Some studies show indeed that when individuals are guaranteed anonymity and privacy they tend to be more cooperative in giving in personal information [**?**].

The fact that the answer is publicly available, however, constitutes a threat for the privacy of the individuals. For instance, assume that we are interested in the query "what is the percentage of individuals with a given disease?". The addition of an individual to the database will modify the percentage, and reveal whether the individual has the disease or not.

A common solution to the above problem is to introduce some output perturbation mechanism based on *randomization*: instead of the exact answer to the query, we report a "noisy" answer. Namely, we use some randomized function $\mathcal{K}$ which produces values according to some probability distribution. Of course, depending on the distribution, it may still be possible to guess the value of an individual with a high probability of success, i.e. there may still be a risk of violating privacy. The notion of *differential privacy*, due to Dwork [12, 15, 13, 14], is a proposal to control such a risk. The idea is to say that $\mathcal{K}$ provides $\epsilon$-differential privacy (for some $\epsilon > 0$) if the ratio between the probabilities that two adjacent databases give the same answer is bound by $e^{\epsilon}$, where by "adjacent" we mean that the databases differ for only one individual. Often we will abbreviate "$\epsilon$-differential privacy" as $\epsilon$-d.p.

Obviously, the smaller is $\epsilon$, the greater is the privacy protection. In particular, when $\epsilon$ is close to 0 the output of $\mathcal{K}$ is nearly independent from the input (all distributions are almost equal). Unfortunately, such $\mathcal{K}$ is practically useless. The *utility*, i.e. the capability to retrieve accurate answers from the reported ones, is the other important characteristic of $\mathcal{K}$, and it is clear that there is a trade-off between utility and privacy. These two notions, however, are not the complete opposite of each other, because utility concerns the relation between the reported answer and the real answer, while privacy is concerns the relation between the reported answer and the information in the database. This asymmetry makes more interesting the problem of finding a good compromise between the two.

At this point, we would like to remark an intriguing analogy between the area of differential privacy and that of quantitative information flow (QIF), both in the motivations and in the conceptual framework. At the motivational level, the concern about privacy is akin the concern about information leakage. At the conceptual level, the randomized function $\mathcal{K}$ can be seen as an information-theoretic channel, and the limit case of $\epsilon = 0$, for which the privacy protection is total, corresponds to the case in which the answer to the query does not add any information about the input distribution, and therefore the leakage is 0.

In this notes we recall the notion of differential privacy and its implications, in light of the min-entropy framework developed for QIF. In particular, we explore the relation between $\epsilon$-d.p., leakage, and utility.

The rest of the paper is organized as follows: in the next section we discuss and compare Shannon entropy and Rényi min-entropy, and their interpretation

in terms of attack models. In Section 3 we illustrate the interpretation of systems as channels, from the point of view of leakage, and we review the main results concerning the information-theoretic approaches based on these two notions of entropy. In Section 4 we show that $\epsilon$-differential privacy induces a bound on the leakage of statistical databases. In Section 5 we show that $\epsilon$-differential privacy induces a bound also on their utility, and we present a method to define a randomization mechanism that gives the maximum utility while providing $\epsilon$-differential privacy. In Section 6 we discuss some related work on the relation between differential privacy and QIF.

## 2 Adversary models and entropy notions

In this section we review the two main notions of entropy used in the literature on quantitative information flow, and we discuss the relation with the model of attacker.

In general we consider the kind of threats that in the model of [**?**] are called *brute-force guessing attacks*, which can be summarized as follows: The goal of the adversary is to determine the value of a random variable. He can make a series of queries to an oracle. Each query must have a yes/no answer. In general the adversary is *adaptive*, i.e. he can choose the next query depending on the answer to the previous ones. We assume that the adversary knows the probability distribution.

In the following, $X, Y$ denote two discrete random variables with finite carriers $\mathcal{X} = \{x_1, \ldots, x_n\}$, $\mathcal{Y} = \{y_1, \ldots, y_m\}$, and probability distributions $p_X(\cdot)$, $p_Y(\cdot)$, respectively. We will use $X \wedge Y$ to represent the random variable with carrier $\mathcal{X} \times \mathcal{Y}$ and joint probability distribution $p_{X \wedge Y}(x, y) = p_X(x) \cdot p_{Y|X}(y \mid x)$, where $p_{Y|X}(y \mid x)$ is the conditional probability that $Y = y$ when $X = x$. The notation $X \cdot Y$ will denote the random variable with carrier $\mathcal{X} \times \mathcal{Y}$ and probability distribution defined as product, i.e. $p_{X \cdot Y}(x, y) = p_X(x) \cdot p_Y(y)$. Clearly if $X$ and $Y$ are independent we have $X \wedge Y = X \cdot Y$. We shall omit the subscripts on the probabilities when they are clear from the context. In general, $p_X(\cdot)$ is called *a priori* distribution of $X$, and the conditional probability $p_{X|Y}(\cdot \mid \cdot)$ is called *a posteriori* distribution of $X$ given $Y$. We will also refer to the individual probabilities $p(x)$ and $p(x|y)$ as the a priori an the a posteriori probabilities of $x$.

### 2.1 Shannon entropy

In this section we illustrate a model of adversary that corresponds to Shannon entropy. The ideas discussed here are based on the works [24, 18].

**Definition 1.** *A Shannon adversary is characterized by questions of the form "is $X \in \mathcal{X}'$?", where $\mathcal{X}' \subseteq \mathcal{X}$. Her goal is to determine exactly the value of $X$, and to do it as fast as possible, i.e. with a minimum number of questions.*

Clearly the strategy of the adversary in choosing the sets $\mathcal{X}'$ will have an influence on the number of queries necessary to determine the value of $X$. Intuitively the best strategy is to choose $\mathcal{X}'$ so that its mass probability is as close as possible to that of $\mathcal{X}'' \setminus \mathcal{X}'$, where $\mathcal{X}''$ is the set of values that are currently determined as possible for $X$. The next example illustrates the situation.

*Example 3.* Consider the set $\mathcal{X} = \{a, b, c, d, e, f, g, h\}$, and assume that the probability is distributed as follows:

$$p(a) = p(b) = \frac{1}{4} \qquad p(c) = p(d) = \frac{1}{8} \qquad p(e) = p(f) = p(g) = p(h) = \frac{1}{16}$$

Fig. 2 illustrates a possible strategy of the adversary. The leafs represent the secrets, i.e. the elements of $\mathcal{X}$. Each internal node is labeled with a set $\mathcal{X}' \subseteq \mathcal{X}$, representing the query "is $X \in \mathcal{X}'$?". Depending on the answer to the query, the adversary will continue searching in the left or in the right subtree. In this particular case, at each step the left and the right subtrees have the same probability mass, and as we will see later this means that the strategy is optimal.

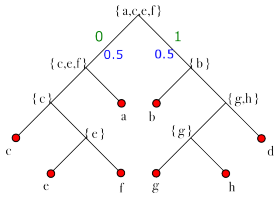



**Fig. 2.** Search space

For a given strategy $\sigma$ of the adversary, let $n_\sigma(x)$ be the number of questions that are needed to determine the value of $X$ when $X = x$. For instance, in the example above, we have $n_\sigma(a) = 2$, and $n_\sigma(e) = 4$. The expected value of $n_\sigma$ is:

$$E_{n_\sigma} \;=\; p(a)\; n_\sigma(a) + \ldots + p(f)\; n_\sigma(f) \;=\; 2 \times \frac{1}{4}2 + 2 \times \frac{1}{8}3 + 4 \times \frac{1}{16}4 \;=\; \frac{11}{4}$$

Note that if we replace $n_\sigma(\cdot)$ by $\log_2 p(\cdot)$ in the above definition of $E_{n_\sigma}$, we obtain the formula for Shannon entropy (cf. Definition 2).

Coming back to Fig. 2, we wish remark that each element $x \in \mathcal{X}$ is uniquely associated to the path from the root to the leaf associated to $x$. If we label the arcs of the tree with 0 (left) and 1 (right), then each $x \in \mathcal{X}$ can be represented by the binary string relative to the path. The problem of minimizing the expected value of $n_\sigma$ corresponds therefore to that of finding the *optimal coding*, i.e. the coding that minimize the expected number of bits needed to represent the elements of $\mathcal{X}$. Shannon entropy is also related to the optimal coding, and more precisely it corresponds to the expected length of the code.

**Definition 2 ([26]).** *The Shannon entropy of a random variable $X$ is defined as*

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x)$$

The minimum value $H(X) = 0$ is obtained when $p(\cdot)$ is concentrated on a single value (i.e. when $p(\cdot)$ is a Dirac measure [1]). The maximum value $H(X) = \log_2 |\mathcal{X}|$ is obtained when $p(\cdot)$ is the uniform distribution. Usually the base of the logarithm is set to be 2 and, correspondingly, the entropy is measured in *bits*.

**Proposition 1.** *$H(X)$ is a lower bound for the expected value of $n_\sigma(\cdot)$, with respect to all possible strategies $\sigma$.*

Note that the lover bound mentioned in the above proposition (i.e. the entropy) is actually achieved in Example 3. Indeed, a simple calculation shows that $H(X) = \frac{11}{4}$. In general the lower bound $H(X)$ can be achieved when for each $x \in \mathcal{X}$ we have that $p(x)$ is a power of 2, and we can organize the questions as a tree, with the characteristic that each node (question) splits the probability mass in two.

We end this section with the definition of conditional entropy.

**Definition 3.** *The* conditional entropy *of $X$ given $Y$ is*

$$H(X \mid Y) = \sum_{y \in \mathcal{Y}} p(y)\ H(X \mid Y = y) \tag{2}$$

*where*

$$H(X \mid Y = y) = -\sum_{x \in \mathcal{X}} p(x|y) \log_2 p(x|y)$$

It is possible to prove that $0 \leq H(X \mid Y) \leq H(X)$. The minimum value, 0, is obtained when $X$ is completely determined by $Y$. The maximum value, $H(X)$, is obtained when $Y$ reveals no information about $X$, i.e. when $X$ and $Y$ are independent.

### 2.2 Rényi min-entropy

In this section we illustrate a model of adversary that corresponds to Rényi min-entropy (or simply min-entropy). This material is based on [27].

**Definition 4.** *In the one-try model the adversary is allowed to ask exactly one question, which must be of the form: "is $X = x$?". Her goal is to maximize the probability of guessing the right element in just one single try.*

Of course, the best strategy for the adversary consists in choosing the $x$ with the maximum probability. Therefore the measure of success of this kind of adversary is $\max_{x \in \mathcal{X}} p(x)$.

We discuss now how the probability of success is related to the Rényi min-entropy. Let us first give a quick overview of the context in which this concept originated.

[1] A *Dirac measure* or *point mass* is a distribution $\delta_w(x)$ that has the value 1 on the point $x = w$ and 0 otherwise.

In [25], Rényi introduced a one-parameter family of entropy measures, intended as a generalization of Shannon entropy. The Rényi entropy of order $\alpha$ ($\alpha > 0$, $\alpha \neq 1$) of a random variable $X$ is defined as

$$H_\alpha(X) \;=\; \frac{1}{1-\alpha} \log_2 \sum_{x \in \mathcal{X}} p(x)^\alpha$$

Rényi's motivations were of axiomatic nature: Shannon entropy satisfies four axioms, namely symmetry, continuity, value 1 on the Bernoulli uniform distribution, and the chain rule[2]:

$$H(X \wedge Y) \;=\; H(X \,|\, Y) + H(Y) \tag{3}$$

Shannon entropy is also the *only* function that satisfies those axioms. However, if we replace (3) with a weaker property representing the additivity of entropy for independent distributions:

$$H(X \cdot Y) \;=\; H(X) + H(Y)$$

then there are more functions satisfying the axioms, among which all those of the Rényi's family.

Shannon entropy is obtained by taking the limit of $H_\alpha$ as $\alpha$ approaches 1. In fact we can easily prove, using l'Hôpital's rule, that

$$H_1(X) \;\stackrel{\text{def}}{=}\; \lim_{\alpha \to 1} H_\alpha(X) \;=\; -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x)$$

We are particularly interested in the limit of $H_\alpha$ as $\alpha$ approaches $\infty$. This is called *min-entropy*. It can be proven that

$$H_\infty(X) \;\stackrel{\text{def}}{=}\; \lim_{\alpha \to \infty} H_\alpha(X) \;=\; -\log_2 \max_{x \in \mathcal{X}} p(x)$$

which gives the connection with the model of adversary described in Definition 4.

As for the $\alpha$-generalization of the conditional entropy, Rényi did not define it, and there is no agreement on what it should be. Various researchers, including Cachin [5], have considered the following definition, based on (2):

$$H_\alpha^{\text{Cachin}}(X \mid Y) = \sum_{y \in \mathcal{Y}} p(y)\; H_\alpha(X \mid Y = y)$$

which, as $\alpha \to \infty$, becomes

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

[2] The original axiom, called the grouping axiom, does not mention the conditional entropy. However it corresponds to the chain rule if the conditional entropy is defined as in (2).

An alternative proposal for $H_\infty(\cdot \mid \cdot)$ came from [**?**,11], and was again advocated by Smith in his seminal paper [27]:

$$H_\infty^{\text{Smith}}(X \mid Y) \;=\; -\log_2 \textstyle\sum_{y\in\mathcal{Y}} \max_{x\in\mathcal{X}} \; p(x,y) \tag{4}$$

The most interesting version of $H_\infty(X \mid Y)$, in terms of security, seems to be that of Smith: Indeed, $\sum_{y\in\mathcal{Y}} \max_{x\in\mathcal{X}} \; p(x,y)$ represents the expected value of the a posteriori probability of the adversary's success. The complement of this expected value is also known as *probability of error* or *Bayes risk*, and has been used as a measure of information flow also by other authors [7]. Observing $Y$ decreases the probability of error, and consequently we can prove formally that $H_\infty^{\text{Smith}}(X \mid Y) \leq H_\infty(X)$, with equality if $X$ and $Y$ are independent. This inequality will ensure that the leakage is always nonnegative, and it is therefore another reason to choose $H_\infty^{\text{Smith}}(\cdot \mid \cdot)$: the alternative $H_\infty^{\text{Cachin}}(\cdot \mid \cdot)$ does not have this feature.

For the sake of simplicity, in the following we will omit the superscript "Smith" and simply write $H_\infty(\cdot \mid \cdot)$.

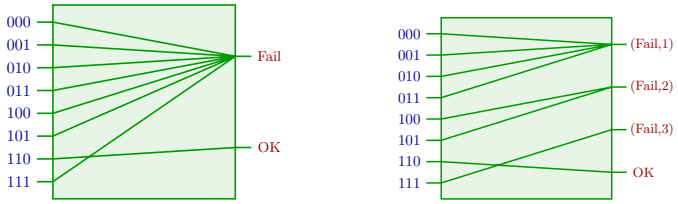
## 3 Information leakage and channels

In this section we consider the interpretation of the leakage in computational systems in terms of channels, and we review some results concerning the Shannon and the min-entropy approaches.

The main idea is that if we focus on the problem of the leakage of secret information trough public observables, then we can regard a computational system as an information-theoretic channel, where the secrets are the inputs, and the observables are the outputs. In the following examples we revisit the ones in the introduction to illustrate this principle.

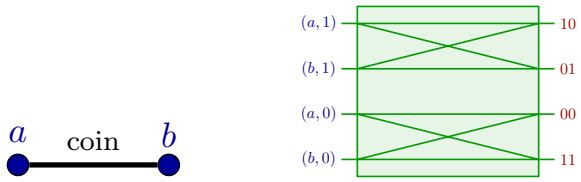
*Example 4.* Consider again the password-checker algorithm defined in Example 1. Assume that the password is 110. Depending on whether the adversary is supposed to be able to detect the number of iterations or not, we have two different channels, as illustrated in the figure below. In the first one, the observable output of the system is either an acceptance of the entered code (OK) or a rejection (Fail). In the second, the observable can be an acceptance (OK), or a rejection together with the number $i$ of iteractions executed (the tuple (Fail,$i$)). Note that the program is deterministic, and this is reflected by the fact that, in both channels, each input corresponds to only one output.

*Example 5.* Consider the dc-nets of Example 2, and consider a particularly simple case: a net with two nodes only, connected by an edge, as illustrated in Fig. 4(a). The channel, illustrated in Fig. 4(b), takes as input the identity of the agent and the bit that she wants to broadcast[3]. On the right-hand side, the each




(a) No iteration count (b) Counting the iterations

**Fig. 3.** Channels associated with the program in Example 1. The password is 110.




(a) A dc-net with two nodes (b) Channel associated with the two-nodes dc-net

**Fig. 4.** A simple dc-net and its channel.

bit of the string represents the declarations of one of the two nodes. Note that this program is not deterministic. The input $(a, 1)$, for instance, can produce two different outcomes: the declaration 10, produced when the coin is 0 (assuming that the first bit from left to right is the declaration of $a$). The other possible declaration, 01, is produced when the coin is 1.

In general, a channel is probabilistic, and it is characterized by its channel matrix.

**Definition 5.** *An information-theoretic channel is a triple $(X, Y, M)$ where $X$ and $Y$ are two random variables representing the input and the output, respectively, and $M$ is a matrix (called the channel matrix) which contains the conditional probabilities $p(y|x)$ for each $x \in \mathcal{X}$ and each $y \in \mathcal{Y}$.*

*Example 6.* The matrices associated to the channels in Figures 3(a) and 3(b) are given in Tables 2(a) and 2(b), respectively. Note that they contain only 0's and 1's. This is characteristic of deterministic channels.

*Example 7.* Consider again Example 2. The channel matrix depends on the coins: whether they are biased, and how much biased they are. Here we consider the simple two-nodes net of Example 5 and two particular situations: the

[3] One may argue that the bit to be broadcasted should not be considered as a secret, and that having it as input will alter the computation of the leakage. This is a valid observation, but there is no harm in representing this bit as input, because we can still remove its contribution to the leakage when we make the calculation.

(a) No iter. c.

| | Fail | OK |
|---|---|---|
| 000 | 1 | 0 |
| 001 | 1 | 0 |
| 010 | 1 | 0 |
| 011 | 1 | 0 |
| 100 | 1 | 0 |
| 101 | 1 | 0 |
| 110 | 0 | 1 |
| 111 | 1 | 0 |

(b) Counting the iterations

| | (Fail, 1) | (Fail, 2) | (Fail, 3) | OK |
|---|---|---|---|---|
| 000 | 1 | 0 | 0 | 0 |
| 001 | 1 | 0 | 0 | 0 |
| 010 | 1 | 0 | 0 | 0 |
| 011 | 1 | 0 | 0 | 0 |
| 100 | 0 | 1 | 0 | 0 |
| 101 | 0 | 1 | 0 | 0 |
| 110 | 0 | 0 | 0 | 1 |
| 111 | 0 | 0 | 1 | 0 |

**Table 2.** Channel matrices for the password-checker algorithm

case in which the coins are unbiased (cf. Table 3(a)) and the case in which they are all biased in favor of 0, and more precisely, they give 0 with probability $\frac{2}{3}$ (cf. Table 3(b)).

(a) Fair coin

| | 10 | 01 | 00 | 11 |
|---|---|---|---|---|
| $(a,1)$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 | 0 |
| $(b,1)$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 | 0 |
| $(a,0)$ | 0 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $(b,0)$ | 0 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |

(b) Biased coin ($p(0) = \frac{2}{3}$, $p(1) = \frac{1}{3}$)

| | 10 | 01 | 00 | 11 |
|---|---|---|---|---|
| $(a,1)$ | $\frac{2}{3}$ | $\frac{1}{3}$ | 0 | 0 |
| $(b,1)$ | $\frac{1}{3}$ | $\frac{2}{3}$ | 0 | 0 |
| $(a,0)$ | 0 | 0 | $\frac{2}{3}$ | $\frac{1}{3}$ |
| $(b,0)$ | 0 | 0 | $\frac{2}{3}$ | $\frac{1}{3}$ |

**Table 3.** Channel matrices for the two-node dc-net of Example 5

We consider now the definition of leakage for the two approaches.

### 3.1 Leakage in the Shannon approach

As explained in the introduction, we regard the leakage as the difference between the initial uncertainty and the remaining uncertainty after the observation. Since the uncertainty is represented by the entropy, the leakage can be represented by the well-known (information-theoretic) notion of mutual information, which is the difference between the entropy of the input and the conditional entropy of the input given the output:

**Definition 6.** *The* mutual information *between $A$ and $B$ is defined as*

$$I(X;Y) \;=\; H(X) - H(X \mid Y)$$

*and it measures the amount of information about $X$ that we gain by observing $Y$.*

It can be shown that $I(X;Y) = I(Y;X)$ and $0 \leq I(X;Y) \leq H(X)$. The notion of capacity, defined as

$$C = \max_{p_X(\cdot)} I(X;Y)$$

is often used to represent the worst-case leakage.

We end this section with some relevant properties of the Shannon capacity:

**Proposition 2.** *The capacity of a channel is 0 if and only if all the rows of the matrix are identical. The capacity is maximum (i.e. equal to $H(X)$) if each column has at most one positive value. If $|X| \leq |Y|$, then also the converse holds.*

### 3.2 Leakage in the min-entropy approach

Also in the case of min-entropy the leakage can be expressed using a concept analogous to mutual information[4]:

$$I_\infty(X;Y) = H_\infty(X) - H_\infty(X \mid Y)$$

and analogously for the capacity:

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

This min-entropy variant of the capacity is called *min-capacity*. Only a weak form of Proposition 2 holds for $C_\infty$, namely, only the *if* part.

**Proposition 3.** *The min-capacity of a channel is 0 if all the rows of the matrix are identical. The capacity is maximum (i.e. equal to $H_\infty(X)$) if each column has at most one positive value.*

It has been proven in [4] that $C_\infty$ is always obtained at the uniform distribution (although this is not necessarily the only possible distribution to achieve capacity), and that it is equal to the sum of the maxima of each column in the channel matrix, i.e. $C_\infty = \sum_{b \in \mathcal{B}} \max_{a \in \mathcal{A}} p(b \mid a)$.

## 4 Differential privacy: leakage

In this section we discuss the notion of differential privacy and we show the relation with the notion of leakage. We only consider here the min-entropy notion of leakage.

[4] Some authors, including Smith, prefer avoiding the notation $I_\infty$ and referring to it as "mutual information" because some of the properties of Shannon mutual information are not satisfied in the min-entropy case. In particular, symmetry is not satisfied, i.e. $I_\infty(X;Y) \neq I_\infty(Y;X)$ in general.

Let us start with some definitions. Let *Ind* be a finite set of individuals that may participate in a database and *Val* a finite set of possible values for the attribute of interest of these individuals. In order to capture in a uniform way the presence/absence of an individual in the database, as well as its value, we assume that the set of possible values has a special element *null* representing the absence of the individual in the database. Thus the set of all possible databases is the set $\mathcal{X} = Val^{Ind}$. We will use $u$ and $v$ to denote the cardinalities of *Ind* and *Val*, $|Ind|$ and $|Val|$, respectively. Hence we have that $|\mathcal{X}| = v^u$. A database $x$ can be represented as a $u$-tuple $v_0 v_1 \dots v_{u-1}$ where each $v_i \in Val$ is the value of the corresponding individual. Two databases $x, x'$ are *adjacent* (or *neighbors*), written $x \sim x'$, if they differ for the value of exactly one individual. For instance, for $u = 3$, $v_0 v_1 v_2$ and $v_0 w_1 v_2$, with $w_1 \neq v_1$, are adjacent. The structure $(\mathcal{X}, \sim)$ forms an undirected graph.

Intuitively, differential privacy is based on the idea that a randomized query function provides sufficient protection if the ratio between the probabilities of two adjacent databases to give a certain answer is bound by $e^\epsilon$, for some given $\epsilon > 0$. Formally:

**Definition 7 ([14]).** *A randomized function $\mathcal{K}$ from $\mathcal{X}$ to $\mathcal{Z}$ provides $\epsilon$-differential privacy if for all pairs $x, x' \in \mathcal{X}$, with $x \sim x'$, and all $S \subseteq \mathcal{Z}$, we have that:*

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

The above definition takes into account the possibility that $\mathcal{Z}$ is a continuous domain. In our case, since $\mathcal{Z}$ is finite, the probability distribution is discrete, and we can rewrite the property of $\epsilon$-d.p. more simply as (using the notation of conditional probabilities, and considering both quotients):

$$\frac{1}{e^\epsilon} \leq \frac{Pr[Z = z | X = x]}{Pr[Z = z | X = x']} \leq e^\epsilon \qquad \text{for all } x, x' \in \mathcal{X} \text{ with } x \sim x', \text{ and all } z \in \mathcal{Z}$$

where $X$ and $Z$ represent the random variables associated to $\mathcal{X}$ and $\mathcal{Z}$, respectively.

### 4.1 Graph symmetries

In this section we explore some classes of graphs that allow us to derive a strict correspondence between $\epsilon$-d.p. and the a posteriori entropy of the input.

Let us first recall some basic notions. 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. The *diameter* of $G$ is the maximum distance between any two vertices in $\mathcal{V}$. The degree of a vertex is the number of edges incident to it. $G$ is called *regular* if every vertex has the same degree. A regular graph with vertices of degree $k$ is called a $k$-regular graph. An automorphism of $G$ is a permutation $\sigma$ of the vertex set $\mathcal{X}$, such that for any pair of vertices $x, x'$, if $x \sim x'$, then $\sigma(x) \sim \sigma(x')$. If $\sigma$ is an automorphism, and $v$ a vertex, the orbit of $v$ under $\sigma$ is the set $\{v, \sigma(v), \dots, \sigma^{k-1}(v)\}$ where $k$ is the smallest positive integer such that $\sigma^k(v) = v$. Clearly, the orbits of the vertices under $\sigma$ define a partition of $\mathcal{V}$.

The following two definitions introduce the classes of graphs that we are interested in. The first class is well known in literature.

**Definition 8.** *Given a graph $G = (\mathcal{V}, \sim)$, we say that $G$ is distance-regular if there exist integers $b_i, c_i, i = 0, ..., d$ such that for any two vertices $v, w$ in $\mathcal{V}$ with distance $i = d(v, w)$, there are exactly $c_i$ neighbors of $w$ in $G_{i-1}(x)$ and $b_i$ neighbors of $v$ in $G_{i+1}(x)$, where $G_i(x)$ is the set of vertices $y$ of $G$ with $d(x, y) = i$.*

The next class is a variant of the VT (vertex-transitive) class:

**Definition 9.** *A graph $G = (\mathcal{V}, \sim)$ is $VT^+$ (vertex-transitive +) if there are $n$ automorphisms $\sigma_0$, $\sigma_1$, $\ldots \sigma_{n-1}$, where $n = |\mathcal{V}|$, such that, for every vertex $v \in \mathcal{V}$, we have that $\{\sigma_i(v) \mid 0 \leq i \leq n-1\} = \mathcal{V}$.*

In particular, the graphs for which there exists an automorphism $\sigma$ which induces only one orbit are $\mathrm{VT}^+$: in fact it is sufficient to define $\sigma_i = \sigma^i$ for all $i$ from 0 to $n - 1$.

From graph theory we know that neither of the two classes subsumes the other. They have however a non-empty intersection, which contains in particular all the structures of the form $(\mathcal{V} \dashv\updownarrow^{Ind}, \sim)$, i.e. the database domains.

**Proposition 4.** *The structure $(\mathcal{X}, \sim) = (\mathcal{V} \dashv\updownarrow^{Ind}, \sim)$ is both a distance-regular graph and a $VT^+$ graph.*

The two symmetry classes defined above, distance-regular and $\mathrm{VT}^+$, will be used in the next section to transform a generic channel matrix into a matrix with a symmetric structure, while preserving the a posteriori min entropy and the $\epsilon$-d.p.. This is the core of our technique to establish the relation between differential privacy and quantitive information flow, depending on the structure induced by the database adjacency relation.

### 4.2 Deriving the relation between differential privacy and QIF on the basis of the graph structure

This section discusses a general technique for determining the relation between $\epsilon$-differential privacy and min-entropy leakage, and between $\epsilon$-differential privacy and utility, depending on the graph structure induced by $\sim$ and $f$. The idea is to use the symmetries of the graph structure to transform the channel matrix into an equivalent matrix with certain regularities, which allow to establish the link between $\epsilon$-differential privacy and the a posteriori min entropy.

Let us illustrate briefly this transformation. Consider a channel whose matrix $M$ has at least as many columns as rows. 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 $\mathrm{VT}^+$, we transform $M'$ into a matrix $M''$ whose diagonal elements are all the same, and coincide with the maximum element of $M''$, which we denote here by $\max^{M''}$.

We are now going to present formally our the technique. Let us first fix some notation: In the rest of this section we consider channels with input $A$ and output $B$, with carriers $\mathcal{A}$ and $\mathcal{B}$ respectively, and we assume that the probability distribution of $A$ is uniform. This is not a restriction for our bounds on the leakage: as seen in the previous section, the maximum min-entropy leakage is achieved in the uniform input distribution and, therefore, any bound for the uniform input 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.

Furthermore, we assume that $|\mathcal{A}| = n \leq |\mathcal{B}| = m$. We also assume 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 will also write $i \sim h$ when $i$ and $h$ are associated to adjacent elements of $\mathcal{A}$, and we will write $d(i, h)$ to denote the distance between the elements of $\mathcal{A}$ associated to $i$ and $h$.

We note that a channel matrix $M$ provides $\epsilon$-d.p. 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 $H_\infty^M(A|B)$.

**Theorem 1.** *Consider a matrix $M$, and let $r$ be a row of $M$. Assume that $(\mathcal{A}, \sim)$ is either distance-regular or VT$^+$, and that $M$ provides $\epsilon$-d.p. For each distance $d$ from 0 to the diameter of $(\mathcal{A}, \sim)$, let $n_d$ be the number of nodes $j$ at distance $d$ from $r$. Then we have that:*

$$H_\infty^M(A|B) \geq -\log_2 \frac{1}{\sum_d \frac{n_d}{e^{\epsilon\, d}}} \tag{5}$$

Note that this bound is tight, in the sense that we can build a matrix for which (5) holds with equality.

In the next section, we will see how to use this theorem for establishing a bound on the leakage and on the utility.

### 4.3 Application to leakage

As already hinted in the introduction, we can regard $\mathcal{K}$ as a channel with input $X$ and output $Z$. From Proposition 4 we know that $(\mathcal{X}, \sim)$ is both distance-regular and VT$^+$, we can therefore apply Theorem 1. Let us fix a particular database $x \in \mathcal{X}$. The number of databases at distance $d$ from $x$ is

$$n_d = \binom{u}{d} (v-1)^d \tag{6}$$

where $u = |Ind|$ and $v = Val$. In fact, recall that $x$ can be represented as a $u$-tuple with values in $\mathcal{V}\dashv\updownarrow$. We need to select $d$ individuals in the $u$-tuple and

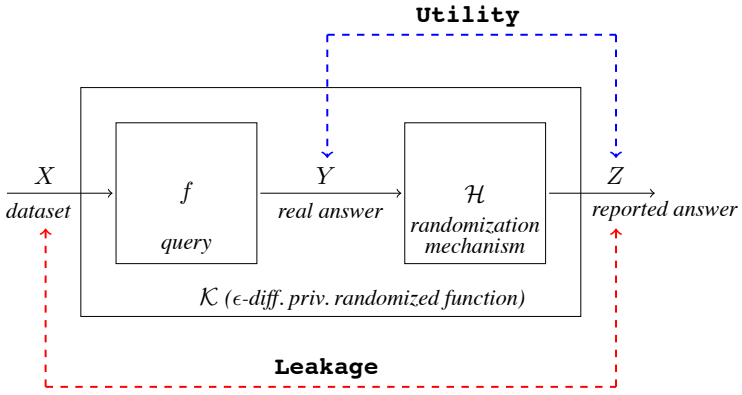



**Fig. 5.** Schema of an oblivious randomized function

then change their values, and each of them can be changed in $v-1$ different ways.

Using the $n_d$ from (6) in Theorem 1 we obtain a binomial expansion in the denominator, namely:

$$H_\infty^M(X|Z) \geq -\log_2 \frac{1}{\sum_{d=0}^{u} \binom{u}{d} (v-1)^d \frac{e^{\epsilon(u-d)}}{e^{\epsilon\, u}}} = -u \log_2 \frac{e^\epsilon}{v-1+e^\epsilon}$$

which gives the following result:

**Theorem 2.** *If $\mathcal{K}$ provides $\epsilon$-d.p., then for the uniform input distribution the information leakage is bound from above as follows:*

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

## 5 Differential privacy: utility

We turn now our attention to the issue of *utility*. We focus on the case in which $\mathcal{K}$ is *oblivious*, which means that it depends only on the (exact) answer to the query, i.e. on the value of $f(x)$, and not on $x$.

An oblivious function can be decomposed in the concatenation of two channels, one representing the function $f$, and the other representing the randomization mechanism $\mathcal{H}$ added as output perturbation. The situation is illustrated in Figure 5.

The standard way to define utility is by means of *guess* and *gain* functions. The functionality of the first is $guess : \mathcal{Z} \to \mathcal{Y}$, and it represents the user's strategy to retrieve the correct answer form the reported one. The functionality

of the latter is $gain : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$. the value $gain(y, y')$ represents the reward for guessing the answer $y$ when the correct answer is $y'$. The utility $\mathcal{U}$ can then be defined as the expected gain:

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

We focus here on the so-called *binary* gain function, which is defined as

$$gain(y, y') = \begin{cases} 1 & \text{if } y = y' \\ 0 & \text{otherwise} \end{cases}$$

This kind of function represents the case in which there is no reason to prefer an answer over the other, except if it is the *right* answer. More precisely, we get a gain if and only if we guess the right answer.

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

**Proposition 5.** *Assume that gain is binary and guess is optimal. Then:*

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

In order to analyze the implications of the $\epsilon$-d.p. requirement on the utility, we need to consider the structure that the adjacency relation induces on $\mathcal{Y}$. Let us define $\sim$ on $\mathcal{Y}$ as follows: $y \sim y'$ if there are $x, x' \in \mathcal{X}$ such that $y = f(x)$, $y' = f(x')$, and $x \sim x'$. Note that $\mathcal{K}$ provides $\epsilon$-d.p. if and only if $\mathcal{H}$ provides $\epsilon$-d.p. (and $\mathcal{H}$ has as inputs only values that are also outputs of $f$).

If $(\mathcal{Y}, \sim)$ is distance-regular or $\mathrm{VT}^+$, then we can apply Theorem 1 to find a bound on the utility. In the following, we assume that the distribution of $Y$ is uniform.

**Theorem 3.** *Consider a randomized mechanism $\mathcal{H}$, and let $y$ be an element of $\mathcal{Y}$. Assume that $(\mathcal{Y}, \sim)$ is either distance-regular or $VT^+$ and that $\mathcal{H}$ provides $\epsilon$-d.p. For each distance $d$ from 0 to the diameter of $(\mathcal{Y}, \sim)$, let $n_d$ be the number of nodes $y'$ at distance $d$ from $y$. Then we have that:*

$$\mathcal{U}(Y, Z) \leq \frac{1}{\sum_{d} \frac{n_d}{e^{\epsilon\, d}}} \tag{7}$$

The above bound is tight, in the sense that (provided $(\mathcal{Y}, \sim)$ is distance-regular or $\mathrm{VT}^+$) we can construct a mechanism $\mathcal{H}$ which satisfies (7) with equality. More precisely, define

$$c = \frac{1}{\sum_{d} \frac{n_d}{e^{\epsilon\, d}}}$$

Then define $\mathcal{H}$ (here identified with its channel matrix for simplicity) as follows:

$$\mathcal{H}_{i,j} = \frac{c}{e^{\epsilon\, d(i,j)}} \tag{8}$$

**Theorem 4.** *Assume $(\mathcal{Y}, \sim)$ is distance-regular or $VT^+$. Then the matrix $\mathcal{H}$ defined in (8) provides $\epsilon$-d.p. and has maximal utility:*

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

Note that we can always define $\mathcal{H}$ as in (8): the matrix so defined will be a legal channel matrix, and it will provide $\epsilon$-d.p.. However, if $(\mathcal{Y}, \sim)$ is neither distance-regular nor $\mathrm{VT}^+$, then the utility of such $\mathcal{H}$ is not necessarily optimal.

We end this section with an example (borrowed from [1]) to illustrate our technique.

*Example 8.* Consider a database with electoral information where each row 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 user voted;
- *Candidate*: the name of the candidate the user voted for.

Consider the query *"What is the city with the greatest number of votes for a given candidate cand?"*. For such a query the binary utility function is the natural choice: only the right city gives some gain, and all wrong answers are equally bad. It is easy to see that every two answers are neighbors, i.e. the graph structure of the answers is a clique.

Let us consider the scenario where the set of cities is $\mathit{Cities} = \{A, B, C, D, E, F\}$ 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 the same candidate. Table 4 shows two alternative mechanisms providing $\epsilon$-differential privacy (with $\epsilon = \log 2$). The first one, $M_1$, is based on the truncated geometric mechanism method used in [16] for counting queries (here extended to the case where every pair of answers is neighbor). The second mechanism, $M_2$, is obtained by applying the definition (8). From Theorem 4 we know that for the uniform input distribution $M_2$ gives optimal utility.

For the uniform input distribution, it is easy to see that $\mathcal{U}(M_1) = 0.2242 < 0.2857 = \mathcal{U}(M_2)$. 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 $\mathcal{U}(M_1) = 0.2412 < 0.2857 = \mathcal{U}(M_2)$. 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 binary[5]. It also works well when

[5] In the metric case the gain function can take into account the proximity of the reported answer to the real one, the idea being that a close answer, even if wrong, is better than a distant one.

(a) $M_1$: truncated geometric mechanism

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

(b) $M_2$: our mechanism

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

**Table 4.** Mechanisms for the city with higher number of votes for candidate *cand*

$(\mathcal{Y}, \sim)$ has low connectivity, in particular in the cases of a ring and of a line. But in this example, we are not in these cases, because we are considering *binary gain functions* and *high connectivity*.

## 6 Related work

As far as we know, the first work to investigate the relation between differential privacy and information-theoretic leakage *for an individual* was [2]. In this work, a channel is relative to a given database $x$, and the channel inputs are all possible databases adjacent to $x$. Two bounds on leakage were presented, one for teh Rényi min entropy, and one for Shannon entropy.

Barthe and Köpf [3] were the first to investigate the (more challenging) connection between differential privacy and the Rényi min-entropy leakage *for the entire universe of possible databases*. They consider the "end-to-end differentially private mechanisms", which correspond to what we call $\mathcal{K}$ in our paper, and propose, like we do, to interpret them as information-theoretic channels. They provide a bound for the leakage, but point out that it is not tight in general, and show that there cannot be a domain-independent bound, by proving that for any number of individual $u$ the optimal bound must be at least a certain expression $f(u, \epsilon)$. Finally, they show that the question of providing optimal upper bounds for the leakage of $\epsilon$-differentially private randomized functions in terms of rational functions of $\epsilon$ is decidable, and leave the actual function as an open question. In our work we used rather different techniques and found (independently) the same function $f(u, \epsilon)$ (the bound in Theorem 1), but we actually proved that $f(u, \epsilon)$ is the optimal bound[6]. Another difference is that [3] captures the case in which the focus of differential privacy is on hiding *participation* of individuals in a database. In our work, we consider both the participation and the *values* of the participants.

Clarkson and Schneider also considered differential privacy as a case study of their proposal for quantification of integrity [10]. There, the authors analyze

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

database privacy conditions from the literature (such as differential privacy, $k$-anonymity, and $l$-diversity) using their framework for utility quantification. In particular, they study the relationship between differential privacy and a notion of leakage (which is different from ours - in particular their definition is based on Shannon entropy) and they provide 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 statical analysis of the information leaked by the query. However [17] did not attempt to relate information leakage to differential privacy.

In [16] the authors aim at obtaining optimal-utility randomization mechanisms while preserving differential privacy. The authors propose 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 (defined as a loss function) the users of $M$ may have. They prove that the geometric mechanism is optimal in the particular case of counting queries. Our results do not restrict to counting queries, but on the other hand we only consider the case of binary loss function.

## References


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. Technical report, 2011. `http://hal.inria.fr/inria-00580122/en/`.
2. Mário S. Alvim, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catuscia Palamidessi. Differential privacy versus quantitative information flow. Technical report, 2010.
3. Gilles Barthe and Boris Köpf. Information-theoretic bounds for differentially private mechanisms. In *Proc. of CSF*, 2011. To appear.
4. Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Quantitative notions of leakage for one-try attacks. In *Proc. of MFPS*, volume 249 of *ENTCS*, pages 75–91. Elsevier, 2009.
5. Christian Cachin. *Entropy Measures and Unconditional Security in Cryptography*. PhD thesis, 1997.
6. Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. Anonymity protocols as noisy channels. *Inf. and Comp.*, 206(2–4):378–401, 2008.
7. Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. On the Bayes risk in information-hiding protocols. *J. of Comp. Security*, 16(5):531–571, 2008.
8. David Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1:65–75, 1988.
9. David Clark, Sebastian Hunt, and Pasquale Malacaria. Quantitative information flow, relations and polymorphic types. *J. of Logic and Computation*, 18(2):181–199, 2005.
10. M. R. Clarkson and F. B. Schneider. Quantification of integrity, 2011. Tech. Rep.. `http://hdl.handle.net/1813/22012`.

11. Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM J. Comput*, 38(1):97–139, 2008.
12. Cynthia Dwork. Differential privacy. In *Automata, Languages and Programming, 33rd Int. Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proc., Part II*, volume 4052 of *LNCS*, pages 1–12. Springer, 2006.
13. Cynthia Dwork. Differential privacy in new settings. In *Proc. of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2010, Austin, Texas, USA, January 17-19, 2010*, pages 174–183. SIAM, 2010.
14. Cynthia Dwork. A firm foundation for private data analysis. *Communications of the ACM*, 54(1):86–96, 2011.
15. Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In *Proc. of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 371–380. ACM, 2009.
16. Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In *Proc. of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 351–360. ACM, 2009.
17. Jonathan Heusser and Pasquale Malacaria. Applied quantitative information flow and statistical databases. In *Proc. of the Int. Workshop on Formal Aspects in Security and Trust*, volume 5983 of *LNCS*, pages 96–110. Springer, 2009.
18. Boris Köpf and David A. Basin. An information-theoretic model for adaptive side-channel attacks. In *Proc. of CCS*, pages 286–296. ACM, 2007.
19. Pasquale Malacaria. Assessing security threats of looping constructs. In *Proc. of POPL*, pages 225–235. ACM, 2007.
20. Pasquale Malacaria and Han Chen. Lagrange multipliers and maximum information leakage in different observational models. In *Proc. of PLAS*, pages 135–146. ACM, 2008.
21. Massey. Guessing and entropy. In *Proc. of ISIT*, page 204. IEEE, 1994.
22. Ira S. Moskowitz, Richard E. Newman, Daniel P. Crepeau, and Allen R. Miller. Covert channels and anonymizing networks. In *Proc. of PES*, pages 79–88. ACM, 2003.
23. Ira S. Moskowitz, Richard E. Newman, and Paul F. Syverson. Quasi-anonymous channels. In *Proc. of CNIS*, pages 126–131. IASTED, 2003.
24. Pliam. On the incomparability of entropy and marginal guesswork in brute-force attacks. In *Proc. of INDOCRYPT*, number 1977 in LNCS, pages 67–79. Springer-Verlag, 2000.
25. Alfréd Rényi. On Measures of Entropy and Information. In *Proc. of the 4th Berkeley Symposium on Mathematics, Statistics, and Probability*, pages 547–561, 1961.
26. Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 625–56, 1948.
27. Geoffrey Smith. On the foundations of quantitative information flow. In *Proc. of FOSSACS*, volume 5504 of *LNCS*, pages 288–302. Springer, 2009.
28. Ye Zhu and Riccardo Bettati. Anonymity vs. information leakage in anonymity systems. In *Proc. of ICDCS*, pages 514–524. IEEE, 2005.