From t-Closeness-Like Privacy to Postrandomization via Information Theory

David Rebollo-Monedero, Jordi Forné, and Josep Domingo-Ferrer

Abstract—t-Closeness is a privacy model recently defined for data anonymization. A data set is said to satisfy t-closeness if, for each group of records sharing a combination of key attributes, the distance between the distribution of a confidential attribute in the group and the distribution of the attribute in the entire data set is no more than a threshold t. Here, we define a privacy measure in terms of information theory, similar to t-closeness. Then, we use the tools of that theory to show that our privacy measure can be achieved by the postrandomization method (PRAM) for masking in the discrete case, and by a form of noise addition in the general case.

Index Terms—t-Closeness, microdata anonymization, information theory, rate-distortion theory, PRAM, noise addition.

1 Introduction

A microdata set is a data set whose records carry information on invidual respondents, like people or enterprises. The attributes in a microdata set can be classified as follows:

- *Identifiers*. These are attributes that *unambiguously* identify the respondent. Examples are passport number, social security number, full name, etc. Since our objective is to prevent confidential information from being linked to specific respondents, we shall assume in what follows that, in a preprocessing step, identifiers have been removed or encrypted.
- *Key attributes*. Borrowing the definition from [2], [3], key attributes are those that, in combination, can be linked with external information to reidentify (some of) the respondents to whom (some of) the records in the microdata set refer. Examples are job, address, age, gender, etc. Unlike identifiers, key attributes cannot be removed, because any attribute is potentially a key attribute.
- *Confidential outcome attributes*. These are attributes which contain sensitive information on the respondent. Examples are salary, religion, political affiliation, health condition, etc.

The classification of attributes as key or confidential need not be disjoint or objectively unique. Ultimately, it relies on the specific application the microdata set is intended for.

There are several privacy models to anonymize microdata sets. k-Anonymity [3], [4] is probably the best known. However, it presents several shortcomings which have motivated the appearance of enhanced privacy models reviewed below. t-Closeness [5] is one of those recent proposals. Despite its conceptual appeal, t-closeness lacks computational procedures which allow reaching it with minimum data utility loss.

1.1 Contribution and Plan of this Paper

Here, we define a privacy measure similar to the idea of t-closeness and provide an information-theoretic formulation of the privacy-distortion trade-off problem in microdata anonymization. This is done in such a way that the knowledge body of information theory can be used to find a solution to it. The resulting solution turns out to be the postrandomization (PRAM) masking method [6], [7], [8] in the discrete case and a form of noise addition in the general case.

Section 2 reviews the state of the art in k-anonymity-based privacy models. Mathematical conventions and a brief review of information-theoretic concepts are provided in Section 3. Section 4 gives an information-theoretic formulation of the privacy-distortion trade-off problem, similar to t-closeness. Section 5 contains a theoretical analysis of the solution to this problem. Empirical results are reported in Section 6. Conclusions are drawn in Section 7.

2 Background and Motivation

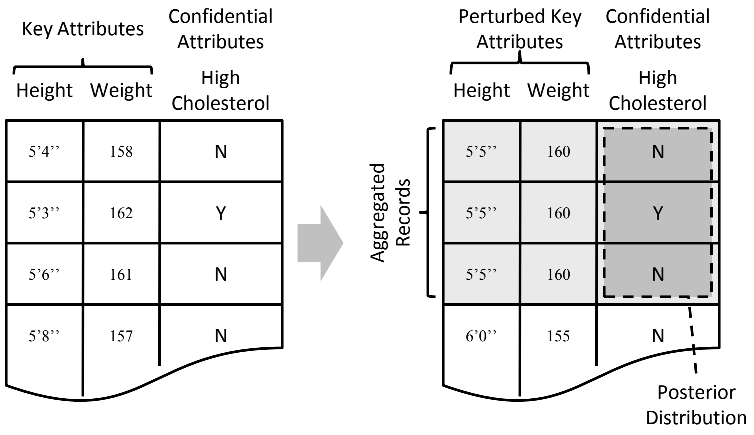
k-Anonymity requires that each combination of key attribute values be shared by at least k records in the data set. To enforce k-anonymity, there are at least two computational procedures: the original approach based

Some parts of this paper (a reduced version of Sections 1 through 4) together with a sketch theoretical analysis for a univariate sensitive attribute were presented at the International Conference on Privacy in Statistical Databases, Istanbul, Turkey, Sep. 2008 [1]. The current multivariate theoretical analysis (Section 5), the experimental work (Section 6), the conclusions (Section 7) and the proofs given in the Appendices are all new work.

- *D. Rebollo-Monedero and J. Forné are with the Department of Telematics Engineering, Technical University of Catalonia, C. Jordi Girona 1-3, E-08034 Barcelona, Catalonia, Spain.*
  E-mail: see http://globus.upc.es/{˜drebollo,˜jforne}.
- *J. Domingo-Ferrer is with the UNESCO Chair in Data Privacy, Department of Computer Engineering and Mathematics, Rovira i Virgili University, Av. Països Catalans 26, E-43007 Tarragona, Catalonia.*
  E-mail: see http://crises-deim.urv.cat/jdomingo.

Manuscript prepared November, 2008; revised April, 2009; second revision July, 2009.

Digital Object Indentifier 10.1109/TKDE.2009.190 1041-4347/$25.00 © 2009 IEEE



Height	Weight	High Cholesterol
5'4''	158	N
5'3''	162	Y
5'6''	161	N
5'8''	157	N

Height	Weight	High Cholesterol
5'5''	160	N
5'5''	160	Y
5'5''	160	N
6'0''	155	N

Fig. 1: Perturbation of key attributes to attain k-anonymity, t-closeness and similar privacy properties.

on generalization and recoding of the key attributes [3], [4] and a microaggregation-based approach described in [9], [10], [11], [12], and illustrated in Fig. 1. While k-anonymity prevents identity disclosure (re-identification is infeasible within a group sharing the same key attribute values), it may fail to protect against (approximate) attribute disclosure: such is the case if the k records sharing a combination of key attribute values also share the value of a confidential attribute. Several enhancements of k-anonymity have been proposed to address the above and other shortcomings. Some of them are mentioned in what follows.

In [13], [14], an evolution of k-anonymity called p-sensitive k-anonymity was presented. Its purpose is to protect against attribute disclosure by requiring that there be at least p different values for each confidential attribute within the records sharing a combination of key attributes. p-Sensitive k-anonymity has the limitation of implicitly assuming that each confidential attribute takes values uniformly over its domain, that is, that the frequencies of the various values of a confidential attribute are similar. When this is not the case, achieving p-sensitive k-anonymity may cause a huge data utility loss.

Like p-sensitive k-anonymity, l-diversity [15] was defined with the aim of solving the attribute disclosure problem that can arise with k-anonymity. A data set is said to satisfy l-diversity if, for each group of records sharing a combination of key attributes, there are at least l "well-represented" values for each confidential attribute. Depending on the definition of "well-represented", l-diversity can reduce to p-sensitive k-anonymity or be a bit more complex. However, it shares with the latter the problem of huge data utility loss. Also, it is insufficient to prevent attribute disclosure, because at least the following two attacks are conceivable:

- *Skewness attack*. If, within a group of records sharing a combination of key attributes, the distribution of the confidential attribute is very different from its distribution in the overall data set, then an intruder linking a specific respondent to that group may learn confidential information (e.g., imagine that the proportion of respondents with AIDS within the group is much higher than in the overall data set).
- *Similarity attack*. If values of a confidential attribute within a group are l-diverse but semantically similar (e.g., similar diseases or similar salaries), attribute disclosure also takes place.

t-Closeness [5] tries to overcome the above attacks. A microdata set is said to satisfy t-closeness if, for each combination of key attributes, the distance between the distribution of the confidential attributes in the group and the distribution of the attributes in the whole data set is no more than a threshold t. t-Closeness can be argued to protect against skewness and similarity (see [16] for a more detailed analysis):

- To the extent to which the within-group distribution of confidential attributes resembles the distribution of those attributes for the entire dataset, skewness attacks will be thwarted.
- Again, since the within-group distribution of confidential attributes mimics the distribution of those attributes over the entire dataset, no semantic similarity can occur within a group that does not occur in the entire dataset. (Of course, within-group similarity cannot be avoided if all patients in a data set have similar diseases.)

The main limitation of the original t-closeness paper is that no computational procedure to reach t-closeness was specified. This is what we address in the remainder of this paper by leaning on the framework of information theory. Throughout Section 4 we provide specific details on additional connections between our work and the literature, from a more technical, information-theoretic perspective.

3 Mathematical Conventions and Information-Theoretic Preliminaries

Throughout the paper, the measurable space in which a random variable (r.v.) takes on values will be called an *alphabet*. We shall follow the convention of using uppercase letters for r.v.'s, and lowercase letters for particular values they take on. Probability density functions (PDFs) and probability mass functions (PMFs) are denoted by p, subindexed by the corresponding r.v. in case of ambiguity risk. For example, both $p_X(x)$ and $p(x)$ denote the value of the function p_X at x, which aids in writing more concise equations. Informally, we occasionally refer to the function p as $p(x)$. Similarly, we use the notations $p_{X|Y}$ and $p(x|y)$ equivalently.

We adopt the same notation for information-theoretic quantities used in [17]. Specifically, the symbol H will denote entropy, h differential entropy, I mutual information, and D relative entropy or Kullback-Leibler (KL) divergence. We briefly recall those concepts for the reader not intimately familiar with information theory:

- The *entropy* $\mathrm{H}(X)$ of a discrete r.v. X with PMF p is a measure of its uncertainty, and it is defined as
$$\mathrm{H}(X) = -\,\mathrm{E}\,\log p(X) = -\sum_x p(x)\,\log p(x).$$
If X is a continuous r.v. instead, say distributed in $\mathbb{R}^k$ and with PDF p, the analogous measure is the *differential entropy*
$$\mathrm{h}(X) = -\,\mathrm{E}\,\log p(X) = -\int_{\mathbb{R}^k} p(x)\,\log p(x)\,\mathrm{d}x.$$
- The *conditional entropy* of a r.v. X given a r.v. Y is the entropy of X conditioned on each value y of Y, averaged over all values y. In the discrete case,
$$\begin{aligned}\mathrm{H}(X|Y) &= -\,\mathrm{E}\,\mathrm{E}\,[\log p(X|Y)|Y] = -\,\mathrm{E}\,\log p(X|Y)\\ &= -\sum_y p(y)\sum_x p(x|y)\,\log p(x|y)\\ &= -\sum_{x,y} p(x,y)\,\log p(x|y),\end{aligned}$$
where $p(x|y)$ is the probability of $X = x$ given that $Y = y$. The *conditional differential entropy* of two continuous r.v.'s is defined in an entirely analogous manner, with expectations taking the form of integrals in lieu of summations.
- Let $p(x)$ and $q(x)$ be two probability distributions over the same alphabet. The *KL divergence* $\mathrm{D}(p\|q)$ is a measure of their discrepancy.In the case when p and q are PMFs, it is defined as
$$\mathrm{D}(p\|q) = \mathrm{E}_p \log\frac{p(X)}{q(X)} = \sum_x p(x)\log\frac{p(x)}{q(x)}.$$
When p and q are PDFs, the expectation is written as an integral. The KL divergence might be thought of as a "distance" between distributions, in the sense that $\mathrm{D}(p\|q) \geqslant 0$, with equality if, and only if, $p = q$ (almost surely).
- The *conditional KL divergence* $\mathrm{D}(p(x|y)\|q(x|y))$ is the average, over the conditioning distribution $p(y)$, of the KL divergence between the conditional distributions $p(x|y)$ and $q(x|y)$ (regarded as unconditional distributions for each y). Precisely, in the discrete case,
$$\begin{aligned}\mathrm{D}(p(x|y)\|q(x|y)) &= \mathrm{E}_{\,p(y)}\,\mathrm{E}_{\,p(x|y)}\left[\log\frac{p(X|Y)}{q(X|Y)}\middle|\,Y\right]\\ &= \mathrm{E}_{\,p(x,y)}\,\log\frac{p(X|Y)}{q(X|Y)} = \sum_{x,y} p(x,y)\,\log\frac{p(x|y)}{q(x|y)},\end{aligned}$$
where the joint distribution is taken to be $p(x,y) = p(x|y)\,p(y)$.
- The *mutual information* $\mathrm{I}(X;Y)$ of two r.v.'s X, Y is a measure of the amount of information that one random variable contains about the other, satisfying $\mathrm{I}(X;Y) \geqslant 0$, with equality if, and only if, X and Y are statistically independent. It is defined as the KL divergence between the joint distribution $p(x,y)$ and the independent distribution $p(x)\,p(y)$ generated by the marginal ones:
$$\mathrm{I}(X;Y) = \mathrm{D}(p(x,y)\|p(x)\,p(y)) = \mathrm{E}\,\log\frac{p(X,Y)}{p(X)\,p(Y)}.$$
In the discrete case,
$$\begin{aligned}\mathrm{I}(X;Y) &= \sum_{x,y} p(x,y)\,\log\frac{p(x,y)}{p(x)\,p(y)}\\ &= \mathrm{H}(X) - \mathrm{H}(X|Y) = \mathrm{H}(Y) - \mathrm{H}(Y|X),\end{aligned}$$
that is, the mutual information is the uncertainty reduction on X if Y is observed, and viceversa. An entirely analogous expression is satisfied in the continuous case, with differential entropies and integrals in lieu of entropies and summations.

A *(deterministic) quantizer* is a function that partitions a continuous range of values x, approximating each resulting cell by a value $\hat{x}$ of a discrete r.v. Often, the quantizer map $\hat{x}(x)$ is broken down into two steps. Namely, an assignment to *quantization indices*, usually natural numbers, by means of a function $q(x)$, and a *reconstruction function* $\hat{x}(q)$ mapping indices into values that approximate the original data, so that $\hat{x}(x) = \hat{x}(q(x))$. A *randomized* version generalizes the function $\hat{x}(x)$ by a conditional PMF $p(\hat{x}|x)$, or $p(q|x)$ in terms of indices.

Occasionally, the expectation of a r.v. is written as the letter μ subindexed by its name. Variances and covariances are written as the squared lowercase letter σ^2 and the uppercase letter Σ, respectively, subindexed by the name of the r.v.'s involved. The notation $X \sim \mathcal{N}(\mu, \Sigma)$ will be used to indicate that X is a Gaussian r.v. with mean μ and positive definite covariance Σ.

The set of real $m \times n$ matrices is denoted by $\mathbb{R}^{m\times n}$. $\operatorname{diag}(d_1, \ldots, d_n)$ denotes a diagonal matrix with entries $d_1, \ldots, d_n$. Curly inequality symbols are used to represent matrix definiteness. For example, $A \succcurlyeq 0$ indicates that A is nonnegative definite.

4 Information-Theoretic Formulation of the Privacy-Distortion Trade-Off

Let W and X be jointly distributed r.v.'s in arbitrary alphabets, possibly discrete or continuous. In the problem of database t-closeness described above and depicted in Fig. 1, X represents (the tuple of) key attributes to be perturbed, which could otherwise be used to identify an individual. In the same application, confidential attributes containing sensitive information are denoted by W. Assume that the joint distribution of X and W is known, for instance, an empirical distribution directly drawn from a table, or a parametric statistical model inferred from a subset of records.

4.1 Distortion Criterion

A *distortion measure* $d(x, \hat{x})$ is any measurable, nonnegative, real-valued function representing the distortion between the original data X and a perturbed version $\hat{X}$, the latter also a r.v., commonly but not necessarily in the same alphabet of X. The associated expected distortion $\mathcal{D} = \mathrm{E}\, d(X, \hat{X})$ provides a measure of utility of the perturbed data, in the intuitive sense that low distortion approximately preserves the values of the original data, and their joint statistical properties with respect to any other data of interest, in particular W. For example, if $d(x, \hat{x}) = \|x - \hat{x}\|^2$, then $\mathcal{D}$ is the mean-square error (MSE). Intuitively, a more general form $d(w, x, \hat{x})$ of distortion measure might interestingly widen the range of applications of our framework, possibly at the cost of mathematical tractability.

4.2 Privacy Criterion

Consider now, on the one hand, the distribution p_W of the confidential information W, and on the other, the conditional distribution $p_{W|\hat{X}}$ given the observation of the perturbed attributes $\hat{X}$. In the database k-anonymization problem, whenever the posterior distribution $p_{W|\hat{X}}$ differs from the prior distribution p_W, we have actually gained some information about individuals statistically linked to the perturbed key attributes $\hat{X}$, in contrast to the statistics of the general population. Concordantly, define the *privacy risk* $\mathcal{R}$ as the *conditional* KL divergence D between the posterior and the prior distributions, that is,

$$\begin{aligned}\mathcal{R} &= \mathrm{D}(p_{W|\hat{X}} \| p_W) = \mathrm{E}_{\hat{X}}\, \mathrm{D}(p_{W|\hat{X}}(\cdot|\hat{X}) \| p_W) \\ &= \mathrm{E}_{\hat{X}}\, \mathrm{E}_{W|\hat{X}} \left[\log \frac{p(W|\hat{X})}{p(W)} \,\middle|\, \hat{X} \right] = \mathrm{E} \log \frac{p(W|\hat{X})}{p(W)}. \quad (1)\end{aligned}$$

Section 3 recalls the concept of conditional KL divergence, also explained, for instance, in [17]. A conditional KL divergence is a KL divergence averaged over a conditioning variable. Conceptually, $\mathcal{R}$ is a measure of discrepancy between $p_{W|\hat{X}}$ and p_W, averaged over $\hat{X}$. Technically, p_W is regarded as a degenerate conditional distribution of W given $\hat{X}$ but in fact independent of $\hat{X}$.

A simple manipulation of (1) shows that

$$\begin{aligned}\mathcal{R} = \mathrm{E} \log \frac{p(W|\hat{X})\, p(\hat{X})}{p(W)\, p(\hat{X})} = \mathrm{E} \log \frac{p(W, \hat{X})}{p(W)\, p(\hat{X})} \\ = \mathrm{I}(W; \hat{X}),\end{aligned}$$

in other words, the privacy risk thus defined coincides with the mutual information (see Section 3) of W and $\hat{X}$. Thus,

$$\mathcal{R} = \mathrm{I}(W; \hat{X}) = \mathrm{H}(W) - \mathrm{H}(W|\hat{X}) \quad (2)$$

in the discrete case, and similarly in terms of differential entropies in the continuous case. Two important remarks are in order. Firstly, owing to the symmetry of mutual information, this shows that the privacy risk (1) may be equivalently defined exchanging the roles of W and $\hat{X}$. Secondly, recall that the KL divergence in (1) vanishes if, and only if, the prior and the posterior distributions match (almost surely), which is equivalent to requiring that the mutual information in (2) vanish, in turn equivalent to requiring that W and $\hat{X}$ be statistically independent. Of course, in this extreme case, the utility of the published data, represented by the distribution $p_{W\hat{X}}$, usually by means of the corresponding table, is severely compromised. In the other extreme, leaving the original data undistorted, i.e., $\hat{X} = X$, compromises privacy, because in general $p_{W|X}$ and p_W differ.

4.3 Connections with Other Privacy Criteria

We would like to stress that the use of an information-theoretic quantity for privacy assessment is by no means new. First and foremost, we would like to acknowledge that our privacy measure is tightly related to the measure of t-closeness in [5]. Direct application of the concept of t-closeness to our formulation would lead to define that a distribution satisfies the criterion of t-closeness provided that $\mathrm{D}(p_{W|\hat{X}}(\cdot|\hat{x}) \| p_W) \leqslant t$ for all values $\hat{x}$ of $\hat{X}$. This would naturally suggest measuring privacy risk as the essential supremum (maximum in the discrete case) of $\mathrm{D}(p_{W|\hat{X}}(\cdot|\hat{X}) \| p_W)$, in lieu of the average of (1). Although in our paper we choose to give credit to the idea of t-closeness with regard to the privacy criterion followed in our formulation, the above discussion clarifies that, technically, below the conceptual level, the t-closeness criterion is not quite the same.

Basically, our privacy criterion is an average measure over a divergence. Recall that, by definition, a divergence itself is also an average measure. In the finite-alphabet case, t-closeness is a maximum over divergences, themselves averages. It might still be questionable whether the original t-closeness also succeeds in capturing the disclosure risk for a particular record. A related, more conservative criterion named δ-disclosure privacy is proposed in [18], which measures the maximum difference between the prior and the posterior distributions for each group sharing a common $\hat{x}$. More precisely, in terms of our formulation and still in the finite case, a distribution satisfies this criterion provided that $\max_w \left| \log \frac{p_{W|\hat{X}}(w|\hat{x})}{p_W(w)} \right| < \delta$ for all values $\hat{x}$ of $\hat{X}$. Simply put, δ-disclosure is a maximum over a maximum.

It is fair to stress that average-case optimization may not address worst cases properly, although the price of worst-case optimization is, in general, a poorer average, *ceteris paribus*. In other words, we must acknowledge that our privacy criterion, in spite of its mathematical tractability, as any criterion based on averages, may not be adequate in all applications [19]. More generally, in spite of its conceptual, information-theoretic appeal, it is important to point out that the adequacy of our formulation relies on the appropriateness of the criteria optimized, which in turn depends on the specific application, on the statistics of the data, on the degree of data utility we are willing to compromise, and last but not least, on the adversarial model and the mechanisms against privacy contemplated. Neither our privacy criterion, nor other widely popular criteria such as k-anonymity in its numerous varieties, are the be-all and end-all of database anonymization [18].

A much earlier connection with the literature can be traced back to the work by Shannon in 1949 [20]. Note that $\mathrm{I}(W;\hat{X})$ and $\mathrm{H}(W|\hat{X})$ in (2) are equivalent minimization objectives in the design of $\hat{X}$, under the assumption that W and therefore $\mathrm{H}(W)$ are given. Shannon introduced the concept of *equivocation* as the conditional entropy of a private message given an observed cryptogram, later used in the formulation of the problem of the wiretap channel [21] as a measure of confidentiality. We can also trace back to the fifties the information-theoretic interpretation of the divergence between a prior and a posterior distribution, named *(average) information gain* in some statistical fields [22], [23].

In addition to the work already cited, [7], [24], [25] already used Shannon entropy as a measure of information loss, pointing out limitations affecting specific applications. We would like to stress out that we use a KL divergence as a measure of *information disclosure* (rather than loss), consistently with the equivalence between the case when $p_{W|\hat{X}} = p_W$ and the complete absence of privacy risk. On the other hand, the flexibility in our definition of distortion measure as a measure of *information loss* may enable us to preserve the statistical properties of the perturbed data to an arbitrary degree, possibly with respect to any other data of interest. Of course, the choice of distortion measure should ultimately rely on each particular application.

4.4 Problem Statement

Consequently, we are interested in the trade-off between two contrasting quantities, privacy and distortion, by means of perturbation of the original data. More precisely, consider *randomized perturbation rules* on the original data X, determined by the conditional distribution $p_{\hat{X}|X}$ of the perturbed data $\hat{X}$ given X. In the special case when the alphabets involved are finite, $p_{\hat{X}|X}$ may be regarded as a transition probability matrix, such as the one that appears in the PRAM masking method [6], [7], [8]. Considering randomized rules with only X as input, but not W, formally assumes the conditional independence of $\hat{X}$ and W given X. Two remarks are in order. First, we consider randomized rules because deterministic quantizers (see Section 3) are a particular case, and at this point we may not discard the possibility that more general rules attain a better trade-off. Secondly, we consider rules that affect and depend on X only, but not W, for simplicity. Specifically, implementing and estimating convenient conditional distributions $p_{\hat{X}|WX}$ rather than $p_{\hat{X}|X}$ will usually be more complex, and require large quantities of data to prevent overfitting issues.

To sum up, we are interested in a randomized perturbation minimizing the privacy risk given a distortion constraint (or viceversa). In mathematical terms, we consistently define the *privacy-distortion function* as

$$\mathcal{R}(\mathcal{D}) = \inf_{\substack{p_{\hat{X}|X} \\ \mathrm{E}\ d(X,\hat{X}) \leqslant \mathcal{D}}} \mathrm{I}(W;\hat{X}). \tag{3}$$

We allow ourselves a small abuse of notation and reuse the letter $\mathcal{D}$ for various mathematical flavors of distortion. These include, on the one hand, the previous definition $\mathcal{D} = \mathrm{E}\ d(X,\hat{X})$ in Section 4.1, and on the other, the bound variable acting as an argument of the above function, where merely $\mathrm{E}\ d(X,\hat{X}) \leqslant \mathcal{D}$ for each value of $\mathcal{D}$, not necessarily with equality. For conceptual convenience, we provide an equivalent definition introducing an auxiliary r.v. Q, playing the role of randomized quantization index, a randomized quantizer $p_{Q|X}$, and a reconstruction function $\hat{x}(q)$ (see Section 3):

$$\mathcal{R}(\mathcal{D}) = \inf_{\substack{p_{Q|X},\, \hat{x}(q) \\ \mathrm{E}\ d(X,\hat{X}) \leqslant \mathcal{D}}} \mathrm{I}(W;Q). \tag{4}$$

Proposition 3 in Section A.1 asserts that there is no loss of generality in assuming that Q and $\hat{X}$ are related bijectively, thus $\mathrm{I}(W;Q) = \mathrm{I}(W;\hat{X})$, and that both definitions indeed lead to the same function. The elements involved in the definition of the privacy-distortion function are depicted in Fig. 2.

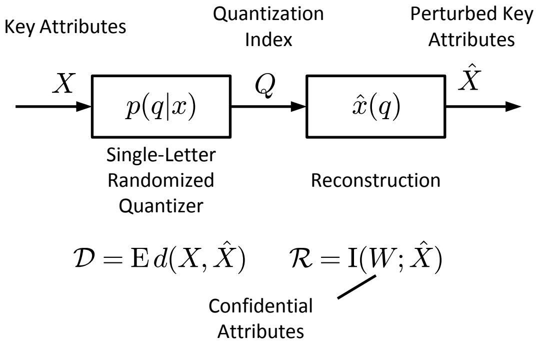


Fig. 2: Information-theoretic formulation of the privacy-distortion problem.

Even though the motivating application for this work is the problem of database t-closeness, it is important to notice that our formulation in principle addresses any applications where perturbative methods for privacy are of interest. Another illustrative application is privacy for location-based services (LBS) [26]. In this scenario, private information such as the user's location (or a sequence thereof) may be modeled by the r.v. X, to be perturbed, and W may represent a user ID. The posterior distribution $p_{\hat{X}|W}$ now becomes the distribution of the user's perturbed location, and the prior distribution $p_{\hat{X}}$, the population's distribution.

4.5 Connections with Rate-Distortion Theory

Perhaps the most attractive aspect of the formulation of the privacy-distortion problem in Section 4.4 is the strong resemblance it bears with the *rate-distortion problem* in the field of information theory. We shall see that our formulation is a generalization of a well-known, extensively studied information-theoretic problem with half a century of maturity. Namely, the problem of lossy compression of source data with a distortion criterion, first proposed by Shannon in 1959 [27].

To emphasize the connection, briefly recall that the simplest version of the problem of lossy data compression, shown in Fig. 3, involves coding of identically distributed (i.i.d.) copies $X_1, X_2, \ldots$ of a generic r.v. X. To this end, an n-letter deterministic quantizer maps blocks of n copies $X_1, \ldots, X_n$ into quantization indices Q in the set $\{1, \ldots, \lfloor 2^{n\mathcal{R}} \rfloor\}$, where $\mathcal{R}$ represents the coding rate in bits needed to represent an index, per sample [17]. An estimation $\hat{X}_1, \ldots, \hat{X}_n$ of the source data vector is recovered to minimize the expected distortion per sample $\mathcal{D} = \frac{1}{n} \sum_i \mathrm{E}\, d(X_i, \hat{X}_i)$, according to some distortion measure $d(x, \hat{x})$. Intuitively, a rate of zero bits may only be achieved in the uninteresting case when no information is conveyed, whereas in the absence of distortion, the rate is maximized. Rate-distortion theory [17] deals with the characterization of the optimal trade-off between the rate $\mathcal{R}$ and the distortion $\mathcal{D}$, allowing codes with arbitrarily large block length n. Accordingly, the *rate-distortion function* is defined as the infimum of the rates of codes satisfying a distortion constraint.

A surprising and fundamental result of rate-distortion theory is that such function, defined in terms of blocks of samples, can be expressed in terms of a single copy or *letter* of the source data vector [17], often more suitable for theoretical analysis. More precisely, the *single-letter characterization of the rate-distortion function* is

$$\mathcal{R}(\mathcal{D}) = \inf_{\substack{p_{\hat{X}|X} \\ \mathrm{E}\, d(X,\hat{X}) \leqslant \mathcal{D}}} \mathrm{I}(X;\hat{X}) = \inf_{\substack{p_{Q|X},\, \hat{x}(q) \\ \mathrm{E}\, d(X,\hat{X}) \leqslant \mathcal{D}}} \mathrm{I}(X;Q), \quad (5)$$

represented in Fig. 4. Aside from the fact that the equivalent problem is expressed in terms of a single letter X rather than n copies, there are two additional differences. First, the quantizer is randomized, and determined by a conditional distribution $p_{Q|X}$. Secondly, the rate is no longer the number of bits required to index quantization cells, or even the lowest achievable rate using an ideal entropy coder, namely the entropy of the quantization index $\mathrm{H}(Q)$. Instead, the rate is a mutual information $\mathcal{R} = \mathrm{I}(X;\hat{X})$.

Interestingly, the single-letter characterization of the rate-distortion function (5) is almost identical to our definition of privacy-distortion function (3), except for the fact that in the latter there is an extra variable W, the confidential attributes, in general different from X, the key attributes. It turns out that some of the information-theoretic results and methods for the rate-distortion problem can be extended, with varying degrees of effort, to the privacy-distortion problem formulated in this work. These extensions are discussed in the next section.

Clearly, the more general privacy-distortion function boils down to Shannon's rate-distortion in the special case when $W = X$. The interpretation behind this case, in light of the formulation of Section 4.4, is that all available attributes are regarded as both key and confidential. In this case, the privacy-distortion function can be computed with the Blahut-Arimoto algorithm [17].

We would like to remark that recent work [28], [29] uses the concept of Shannon's equivocation [20] as a measure of anonymity, similarly to our privacy criterion (2). In addition, it elegantly establishes a relationship between rate-distortion theory on the one hand, and, on the other, the trade-off between throughput and anonymity against traffic analysis in network packet scheduling. It is important to notice that the parallelism in the work cited is for the original rate-distortion function proposed by Shannon (5), expressed in terms of two variables, namely X and $\hat{X}$ in our notation. In our work, however, we extend Shannon's function to three variables (3), namely W, X and $\hat{X}$. This does not only require a substantially more complex theoretical and computational analysis, as we shall see in Sections 5 and 6, but also formulates a substantially more general problem. On a secondary note, our application, randomized database anonymization, is completely different from network traffic anonymization: just note that our application has one more variable W, in general different from X.

5 Theoretical Analysis

This section investigates the privacy-distortion function (3) introduced in Section 4, by means of extending some of the fundamental properties of its information-theoretic analogous, namely the rate-distortion function. Particularly, we confirm that the privacy-distortion function is convex, extend Shannon's lower bound, and analyze the quadratic-Gaussian case in detail.

5.1 Convexity of the Privacy-Distortion Function

The following theorem states that, similarly to the rate-distortion function, the privacy-distortion function is nonincreasing, convex, and therefore continuous in the interior of its domain.

Theorem 1: The privacy-distortion function (3) is nonincreasing and convex.

The proof of the theorem is provided in Section A.2.

Furthermore, the optimization problem determining (3), with $p_{\hat{X}|X}$ as unknown variable, is itself convex. This means that any local minimum is also global, and makes the powerful tools of convex optimization [30] applicable to compute numerically but efficiently the privacy-distortion function. In Section 6, an example of numerical computation will be discussed.

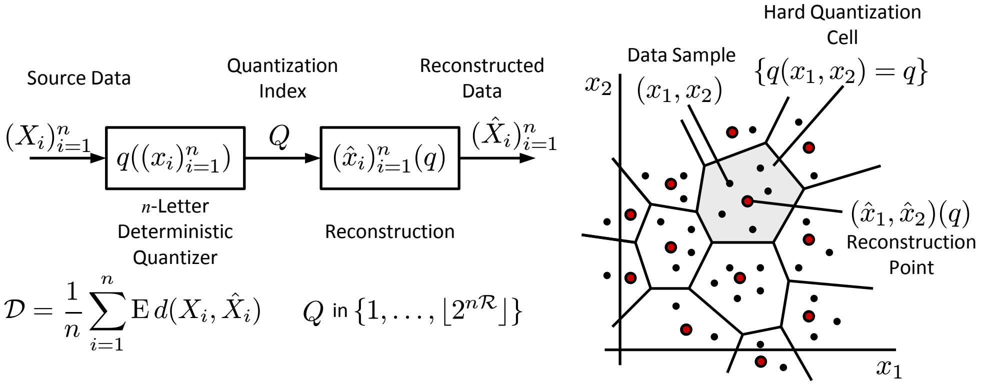


Fig. 3: Information-theoretic formulation of the rate-distortion problem.

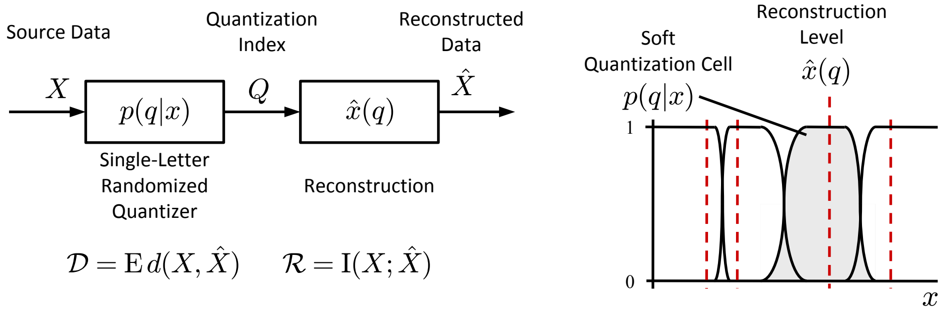


Fig. 4: Single-letter characterization of the rate-distortion problem.

5.2 Quadratic-Gaussian Lower Bound

While a general closed-form expression for privacy-distortion function has not been provided, the Shannon lower bound for the rate-distortion function can be extended to find closed-form lower bounds under certain assumptions. Furthermore, the techniques used to prove this bound may yield an exact closed formula in specific cases. A closed-form upper bound is presented later in this section.

Recall the notational conventions introduced in Section 3. Throughout this section, W and X are r.v.'s taking values in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, and MSE is used as distortion measure, thus $\mathcal{D} = \mathrm{E}\, \|X - \hat{X}\|^2$. The covariance $\Sigma_{W|X}$ of the error of the best (i.e., minimum MSE) linear estimate of W from X is

$$\Sigma_{W|X} = \Sigma_W - \Sigma_{WX}\Sigma_X^{-1}\Sigma_{WX}^{\mathrm{T}}.$$

On account of the next theorem, the following function will be called the *quadratic-Gaussian lower bound* (QGLB):

$$\begin{aligned}\mathcal{R}_{\mathrm{QGLB}}(\mathcal{D}) &= \mathrm{h}(W) - \tfrac{m}{2}\log(2\pi e) \\ &- \max_{\substack{\Sigma \in \mathbb{R}^{n\times n} \\ 0 \preccurlyeq \Sigma \preccurlyeq \Sigma_X \\ \operatorname{tr}\Sigma \leqslant \mathcal{D}}} \tfrac{1}{2}\log\det\left(\Sigma_{W|X} + \Sigma_{WX}\Sigma_X^{-1}\,\Sigma\,\Sigma_X^{-1}\Sigma_{WX}^{\mathrm{T}}\right).\end{aligned} \tag{6}$$

The theorem asserts that the QGLB is indeed a lower bound on the privacy-distortion function (3), for *any joint distribution* of W and X in Euclidean spaces of arbitrary dimension, Gaussian or not. The name is given in recognition of the fact that the bound holds with equality in the Gaussian case, and of the role that properties of Gaussian r.v.'s play in the general proof.

Theorem 2: Provided that W and X are r.v.'s in Euclidean spaces of arbitrary dimension, and that MSE is used as distortion measure, $\mathcal{R}(\mathcal{D}) \geqslant \mathcal{R}_{\mathrm{QGLB}}(\mathcal{D})$ for all $\mathcal{D}$. In the special case when W and X are jointly Gaussian, the bound holds with equality, and the optimal solution Σ in the bound and the optimal solution $\hat{X}$ in the privacy-distortion function are related by $\hat{X} = (I - A)\,X + A\,Z$, where $A = \Sigma\,\Sigma_X^{-1}$ and $Z \sim \mathcal{N}(\mu_X, (I - A)\Sigma_X A^{-\mathrm{T}})$ is independent of X and W.

The proof of the theorem is presented in Appendix B. Fortunately, the matrix optimization problem in the definition of the QGLB (6) is convex. More precisely, it requires the maximization of a convex function subject to linear matrix inequalities [30]. But the major practical advantage over the general expression of the privacy-distortion function (3) is the number of real-valued variables in the new optimization problem (6), given by the size $n \times n$ of the matrix Σ. This number, $n(n-1)/2$ since Σ is symmetric, will be much smaller than the number of variables required in a discretized optimization of the continuous problem posed by (3) (for example, in the simple experiments of Section 6, where $n = 1$, the discretization used involves a matrix representing $p_{\hat{X}|X}$ with 31^2 entries.) Methods to solve these problems numerically are outlined in Appendix C for the QGLB, and in Section 6 for the general privacy-distortion function. Furthermore, Appendix C provides a closed-form, parametric upper bound on the maximum of the optimization problem inherent in the QGLB, and consequently a looser lower bound on the privacy-distortion function. Aside from *not* requiring numerical optimization, this second bound comes with the convenient property of matching the QGLB for small distortion values.

In the special case when W and X are random scalars ($m = n = 1$), direct application of Theorem 2 yields a closed-form expression for the QGLB, and a closed-form solution in the Gaussian case. Define the normalized distortion $d = \frac{\mathcal{D}}{\sigma_X^2}$, where σ_X^2 denotes the variance of X. Let σ_W^2 be the variance of W, and ρ_{WX} the correlation coefficient of W and X. Then,

$$\begin{aligned}\mathcal{R}(\mathcal{D}) &\geqslant \mathcal{R}_{\mathrm{QGLB}}(\mathcal{D}) \\ &= \mathrm{h}(W) - \tfrac{1}{2}\log\left(2\pi e\left(1-(1-d)\rho_{WX}^2\right)\sigma_W^2\right) \quad (7)\end{aligned}$$

for $0 \leqslant d \leqslant 1$ (for $d \geqslant 1$, clearly $\mathcal{R} = 0$).

Provided that W and X are jointly Gaussian random scalars, and that MSE is used as distortion measure, the QGLB (6) is tight:

$$\mathcal{R}(\mathcal{D}) = -\tfrac{1}{2}\log\left(1-(1-d)\rho_{WX}^2\right), \quad (8)$$

with $d = \frac{\mathcal{D}}{\sigma_X^2} \leqslant 1$ as before. The optimal randomized perturbation rule achieving this privacy-distortion performance is represented in Fig. 5. Observe that the perturbed data $\hat{X}$ is a convex combination of the source data X and independent noise, in a way such that the final variance achieves the distortion constraint with equality.

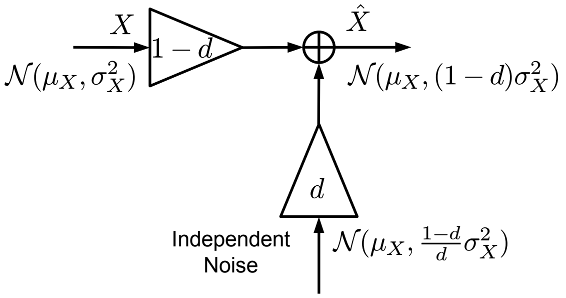

X
$1-d$
$\hat{X}$
$\mathcal{N}(\mu_X, \sigma_X^2)$
$\mathcal{N}(\mu_X, (1-d)\sigma_X^2)$
d
Independent Noise
$\mathcal{N}(\mu_X, \frac{1-d}{d}\sigma_X^2)$

Fig. 5: Optimal randomized perturbation in the quadratic-Gaussian case.

5.3 Mutual-Information Upper Bound

With the same assumptions of multidimensional, Euclidean r.v.'s, and MSE distortion measure, extend the definition of the normalized distortion to $d = \mathcal{D}/\operatorname{tr}\Sigma_X$, and consider the two trivial cases $d = 0$ and $d = 1$. The former case can be achieved with $\hat{X} = X$, yielding $\mathcal{R}(\mathcal{D}) = \mathrm{I}(W;X)$, and the latter with $\hat{X} = \mu_X$, the mean of X, for which $\mathcal{R}(\mathcal{D}) = 0$. Now, for any $0 \leqslant d \leqslant 1$, set $\hat{X} = X$ with probability $1-d$, and $\hat{X} = \mu_X$ with probability d. Convexity properties of the mutual information guarantee that the privacy-distortion performance of this setting cannot lie above the segment connecting the two trivial cases. Since the setting is not necessarily optimal,

$$\mathcal{R}(\mathcal{D}) \leqslant \mathcal{R}_{\mathrm{MIUB}}(\mathcal{D}) = \mathrm{I}(W;X)(1-d). \quad (9)$$

We shall call this bounding function the *mutual-information upper bound* (MIUB). The $p_{\hat{X}|X}$ determined by the combination of the two trivial cases for intermediate values of d may be a simple yet effective way to initialize numerical search methods to compute the privacy-distortion function, as it will be shown in Section 6.

6 Experimental Results

6.1 Computation of the Privacy-Distortion Function

In this section, we illustrate the theoretical analysis of Section 5 with experimental results for a simple, intuitive case. Specifically, W and X are jointly Gaussian random scalars with correlation coefficient ρ (after zero-mean, unit-variance normalization). In terms of the database anonymization problem, W represents sensitive information, and X corresponds to key attributes that can be used to identify specific individuals. These variables could model, for example, the plasma concentration of LDL cholesterol in adults, which is approximately normal, and their weight, respectively. MSE is used as a distortion measure. For convenience $\sigma_X^2 = 1$, thus $d = \mathcal{D}$. Since the privacy-distortion function is convex, minimization of one objective with a constraint on the other is equivalent to the minimization of the Lagrangian cost $\mathcal{C} = \mathcal{D} + \lambda\mathcal{R}$, for some positive multiplier λ. We wish to design randomized perturbation rules $p_{\hat{X}|X}$ minimizing $\mathcal{C}$ for several values of λ, to investigate the feasibility of numerical computation of the privacy-distortion curve, and to verify the theoretical results for the quadratic-Gaussian case of Section 5. As argued in Section 4.4 the perturbation $p_{\hat{X}|X}$ is basically the PRAM masking method in the discrete case, and a form of noise-addition in the continuous case; we take here microaggregation as a noise addition method.

We implement a slight modification of a simple optimization technique, namely the steepest descent algorithm, operating on a sufficiently fine discretization of the variables involved. More precisely, p_{WX} is the joint PMF obtained by discretizing the PDF of W and X, where each variable is quantized with 31 samples in the interval $[-3, 3]$. The starting values for $p_{\hat{X}|X}$ are convex combinations of the extreme cases corresponding to $d = 0$ and $d = 1$, as described in Section 5 when the MIUB (9) was discussed. Only results corresponding to the correlation coefficient $\rho = 0.95$ are shown, for two reasons. First, because of their similarity with results for other values of ρ. Secondly, because for high correlation, the gap between the MIUB (which approximates the performance of the starting solutions) and the QGLB (6) is wider, leading to a more challenging problem.

The definitions of distortion and privacy risk in Section 4 for the finite-alphabet case become

$$\mathcal{D} = \sum_x \sum_{\hat{x}} p(x)p(\hat{x}|x)d(x,\hat{x}),$$

$$\mathcal{R} = \sum_w \sum_{\hat{x}} p(w)p(\hat{x}|w)\ln\frac{p(\hat{x}|w)}{p(\hat{x})}.$$

The conditional independence assumption in the same section enables us to express the PMFs of $\hat{X}$ in the expression for $\mathcal{R}$ as $p(\hat{x}) = \sum_x p(\hat{x}|x)p(x)$ and $p(\hat{x}|w) =$

$\sum_x p(\hat{x}|x)p(x|w)$, in terms of the optimization variables $p(\hat{x}|x)$. Our implementation of the steepest descent algorithm uses the exact gradient with components $\frac{\partial \mathcal{C}}{\partial p(\hat{x}|x)} = \frac{\partial \mathcal{D}}{\partial p(\hat{x}|x)} + \lambda \frac{\partial \mathcal{R}}{\partial p(\hat{x}|x)}$, where $\frac{\partial \mathcal{D}}{\partial p(\hat{x}|x)} = p(x)d(x,\hat{x})$ and

$$\frac{\partial \mathcal{R}}{\partial p(\hat{x}|x)} = p(x)\left(\sum_w p(w|x) \ln p(\hat{x}|w) - \ln p(\hat{x})\right)$$

(after nontrivial simplification).

Two modifications of the standard version of the steepest descent algorithm [30] were applied. First, rather than updating $p_{\hat{X}|X}$ directly according to the negative gradient multiplied by a small factor, we used its projection onto the affine set of conditional probabilities satisfying $\sum_{\hat{x}} p(\hat{x}|x) = 1$ for all x, which in fact gives the steepest descent within that set. Secondly, rather than using a barrier or a Lagrangian function to consider the constraint $p(\hat{x}|x) \geqslant 0$ for all x and $\hat{x}$, after each iteration, we reset possible negative values to 0 and renormalized the probabilities accordingly. This may seem unnecessary since the theoretical analysis in Section 5 gives a strictly feasible solution (i.e., probabilities are strictly positive), and consequently the constraints are inactive. However, the algorithm operates on a discretization of the joint distribution of W and X in a machine with finite precision. The fact is that precision errors in the computation of gradient components corresponding to very low probabilities activated the nonnegativity constraints. Finally, we observed that the ratio between the largest and the smallest eigenvalue of the Hessian matrix was large enough for the algorithm to require a fairly small update factor, 10^{-4}, to prevent significant oscillations.

The privacy-distortion performance of the randomized perturbation rules $p_{\hat{X}|X}$ found by our modification of the steepest descent algorithm is shown in Fig. 6, along with the bounds established in Section 5, namely the QGLB (6) and the MIUB (9). On account of (8), it can be shown that $\lambda = 2\sigma_X^2 \left(1/\rho^2 - 1 + d\right)$. Accordingly, we set λ approximately to 0.72, 1.22 and 1.72, which theoretically corresponds to $d = 0.25, 0.5, 0.75$.

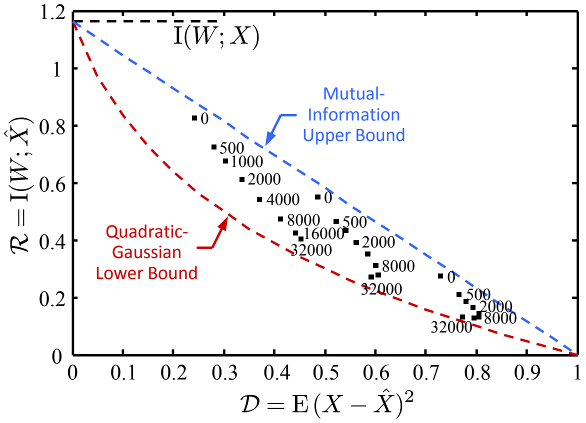


Fig. 6: Privacy-distortion performance of randomized perturbation rules found by a modification of the steepest descent algorithm.

We would like to remark that a completely meaningful interpretation of the values of $\mathcal{D}$ and $\mathcal{R}$ would necessarily lie within the specific application or dataset our formulation is used for. The purpose of this brief experimental section is to illustrate our complex theoretical analysis with quite simple statistics, informally motivated at the beginning of this section by the correlation between cholesterol and weight, but ultimately synthetic. Keep in mind that our framework is not aimed at recommending a particular point in the privacy-distortion plane, but at characterizing a trade-off by means of a curve. Having said that, we attempt to partially interpret the results of Fig. 6, not without a certain degree of abstraction. For instance, the figure shows that for a distortion $\mathcal{D} = \mathrm{E}\,(X - \hat{X})^2 = d = 0.5$ the corresponding privacy risk is roughly $\mathcal{R} = \mathrm{I}(W;\hat{X}) \simeq 0.3$. This means that the randomized perturbation $\hat{X}$ distorts the original data X, leading to an MSE of half of the variance $\sigma_X^2 = 1$ of the unperturbed data. This enormous perturbation in turn reduces the privacy risk, that is, the mutual information, from roughly $\mathrm{I}(W;X) \simeq 1.2$ to $\mathrm{I}(W;\hat{X}) \simeq 0.3$. Recall that mutual information is the amount of information that one variable contains about the other (see Sections 3 and 4.2, and the connection with Shannon's equivocation in Section 4.3). Our theory implies that such ambitious reduction in mutual information is not possible without incurring at least this much data distortion. In addition, the same figure unsurprisingly confirms that perfect privacy $\mathcal{R} = 0$, that is, statistical independence between W and X, is unattainable without an absolutely detrimental impact on the distortion $\mathcal{D} = d = \sigma_X^2 = 1$.

A total of 32000 iterations were computed for each value of λ, at roughly 150 iterations per second on a modern computer[1]. The large number of iterations is consistent with the fact that the Hessian is ill-conditioned and the small updating step size. One would expect that methods based on Newton's technique [30] converge to the optimal solution in less iterations (at the cost of higher computational complexity per iteration), but our goal was to check the performance of one of the simplest optimization algorithms. In all cases, the conditional PMFs found had a performance very close to that described by (8) in Section 5. Their shape, depicted in Fig. 7, roughly resembled the Gaussian shape predicted by the theoretical analysis as the number of iterations increased. The shape of the solution plotted is due to the fact that, intuitively, the steepest descent method moves from a "peaky" initial guess towards the optimum, without quite reaching it. Specifically, Fig. 7 corresponds to $\lambda \simeq 1.22$, was obtained after 32000 iterations, and the number of discretized samples of X and W was increased from 31 to 51. Increasing the number of iterations to 128000 resulted in an experimental solution shaped almost identically to the optimal one, although the one in Fig. 7, corresponding to one fourth of the number of iterations, already achieves values of $\mathcal{C}$ nearly optimal.

[1] Implementation used Matlab R2008a on Windows Vista SP1, on an Intel Core2 Quad Q6600 CPU at 2.4 GHz.

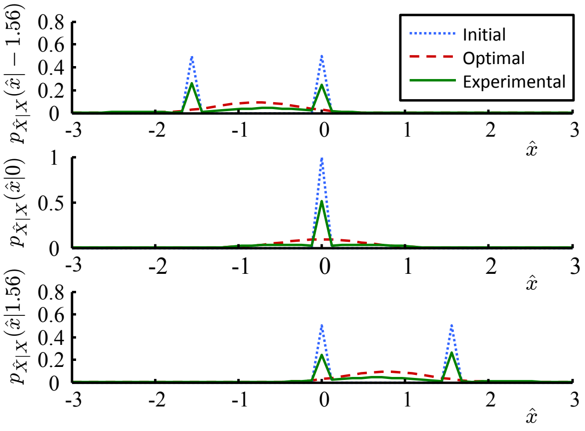


Fig. 7: Shape of initial, optimal, and experimental randomized perturbation rules $p_{\hat{X}|X}$ found by the steepest descent algorithm.

6.2 Performance of MDAV and μ-Approx

Finally, we investigate the privacy-distortion performance of two well-known microaggregation algorithms, namely the maximum distance to average vector (MDAV) algorithm [11], [31] (implemented in the μ-Argus [32] and SDCmicro [33] freeware packages), and the μ-Approx algorithm [12]. The experiment used $n = 2^{15} = 32768$ drawings of W, X, according to the very same Gaussian distribution used in the verification of the QGLB in Fig. 6. The privacy risk $\mathcal{R} = \mathrm{I}(W; \hat{X})$ was estimated from a fine quantization of interval length 0.15 on W and the $\hat{X}$ resulting from the application of each of the k-anonymity algorithms mentioned. The distortion $\mathcal{D} = \mathrm{E}\,(X - \hat{X})^2$ was computed directly from the original samples and the centroids of the aggregation cells.

The results for several values of k/n are shown in Figs. 8 and 9. The most striking finding is the near-optimality of both algorithms in terms of the performance criteria proposed in this work. Intuition suggests, however, that part of their success for this particularly simple data is due to the fact that the key attribute X is a random scalar. The performance of both algorithms would probably deviate more from optimality in case of a multidimensional key attribute (or several key attributes). On the other hand, we would like to mention two reasons why the k-anonymity algorithms do not achieve the QGLB even in this simple setting. First, the perturbation rules considered in the optimization problem inherent in the privacy-distortion function are not constrained to be deterministic. Secondly, the two k-anonymity algorithms investigated operate solely on X, whereas our criterion for privacy risk involves W as well.

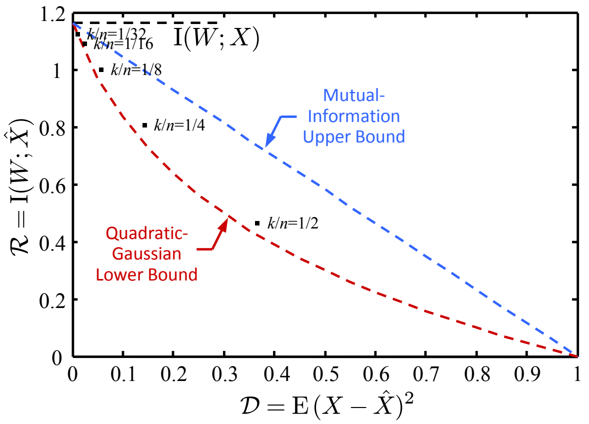


Fig. 8: Privacy-distortion performance of MDAV on scalar data.

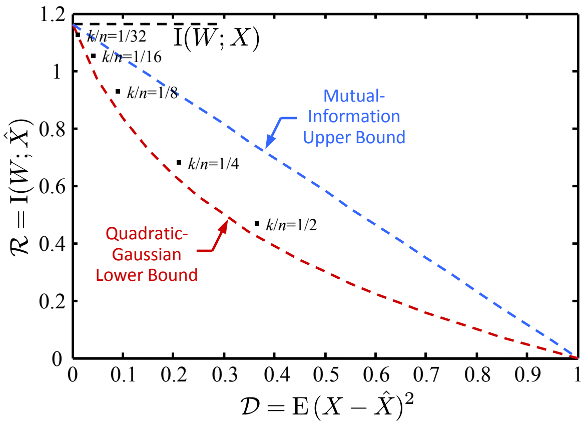


Fig. 9: Privacy-distortion performance of μ-Approx on scalar data.

7 Conclusion

An information-theoretic formulation of the privacy-distortion trade-off in applications such as microdata anonymization and location privacy in location-based services is provided. Inspired by the t-closeness model and Shannon's concept of equivocation, the privacy risk is measured as the mutual information between perturbed key attributes and confidential attributes, equivalent to the conditional KL divergence between posterior and prior distributions. We consider the problem of maximizing privacy (that is, minimizing the above mutual information) while keeping the perturbation of data within a prespecified bound to ensure that data utility is not too damaged.

We establish a strong connection between this privacy-perturbation problem and the rate-distortion problem of information theory and extend of a number of results, including convexity of the privacy-distortion function and the Shannon lower bound.

In principle, it is clear that randomized perturbation rules, being more general, cannot lead to worse performance than deterministic aggregation, such as that used by MDAV and μ-Approx. A privacy-distortion formula is obtained for the quadratic-Gaussian case, proving that the optimal perturbation is in general randomized rather than deterministic, at least in the continuous case. On the other hand, experimental results with discretized statistics show better performance compared to popular deterministic aggregation algorithms. This evidence supports the use of PRAM in the case of attributes with finite alphabets and noise addition in the continuous case.

Appendix A
Fundamental Properties of the Privacy-Distortion Function

A.1 Optimal Reconstruction

The following proposition is a fundamental result on the reconstruction function $\hat{x}(q)$ in the quantizer formulation of the privacy-distortion function.

Proposition 3: Definitions (3) and (4) of the privacy-distortion function are equivalent. Furthermore, without loss of generality, one may assume that Q and $\hat{X}$ are related bijectively, or even that they are equal, in the sense that the assumption does not compromise the optimization problem in the second definition. Finally, in the special case when MSE is used as distortion measure, it may be assumed without loss of generality that $\hat{X} = Q = \mathrm{E}\,[X|Q]$, i.e., $\hat{X} = Q$, where Q is the best (nonlinear) MSE estimate of X from Q.

Proof: For simplicity, we only consider the case when the infimum is in fact a minimum, and when $\mathrm{E}\ d(X, \hat{X})$ can also be minimized. The general proof is similar. Let Q, $\hat{X} = \hat{x}(Q)$ correspond to an optimal solution to the optimization problem in (4). Suppose that Q and $\hat{X}$ are both replaced by a common function of Q minimizing $\mathrm{E}\ d(X, \hat{X})$. In the special case when MSE is used as distortion measure, such function is simply the best MSE estimate $\mathrm{E}\,[X|Q]$ of X from Q. Clearly, this new function of Q will satisfy the distortion constraint. In addition, the data processing inequality [17] guarantees that the mutual information $\mathrm{I}(W; Q)$ cannot increase. This means that the new choice for Q and $\hat{X}$ must also correspond to an optimal solution. ■

A.2 Convexity

We prove that the privacy-distortion function (3) is nonincreasing and convex, as stated in Theorem 1.

Proof: The monotonicity of $\mathcal{R}(\mathcal{D})$ follows immediately from the fact that increasing values of $\mathcal{D}$ relax the minimization constraint in its definition.

To prove convexity, for simplicity, we assume that the infimum is achieved, thereby being a minimum. The general proof is similar. Let $(\mathcal{R}_1, \mathcal{D}_1)$ and $(\mathcal{R}_2, \mathcal{D}_2)$ be pairs on the $\mathcal{R}(\mathcal{D})$ curve, achieved by $p_1(\hat{x}|x)$ and $p_2(\hat{x}|x)$, respectively. Consider the convex combination of randomized rules $p_\lambda = \lambda p_1 + (1-\lambda) p_2$. Since the distortion corresponding to this new rule $\mathcal{D}_\lambda$ is a linear functional of $p_\lambda(\hat{x}|x)$, clearly $\mathcal{D}_\lambda = \lambda \mathcal{D}_1 + (1-\lambda)\mathcal{D}_2$.

Similarly, by construction $W \leftrightarrow X \leftrightarrow \hat{X}_\lambda$, hence $p_\lambda(\hat{x}|w) = \sum_x p_\lambda(\hat{x}|x) p(x|w)$ also depends linearly on the randomized rule, thus

$$p_\lambda(\hat{x}|w) = \lambda p_1(\hat{x}|w) + (1-\lambda) p_2(\hat{x}|w).$$

Recall that mutual information is a convex function of the conditional distribution [17]. Consequently,

$$\mathrm{I}(W; \hat{X}_\lambda) \leqslant \lambda\, \mathrm{I}(W; \hat{X}_1) + (1-\lambda)\, \mathrm{I}(W; \hat{X}_2).$$

Finally, on account of the definition of $\mathcal{R}(\mathcal{D})$,

$$\begin{aligned}\mathcal{R}(\mathcal{D}_\lambda) &\leqslant \mathrm{I}(W; \hat{X}_\lambda)\\ &\leqslant \lambda\, \mathrm{I}(W; \hat{X}_1) + (1-\lambda)\, \mathrm{I}(W; \hat{X}_2)\\ &= \lambda \mathcal{R}(\mathcal{D}_1) + (1-\lambda)\mathcal{R}(\mathcal{D}_2).\end{aligned}$$

■

Appendix B
Quadratic-Gaussian Lower Bound

We prove that the QGLB (6) is in fact a lower bound of the privacy-distortion function (3), as stated in Theorem 2 in Section 5.2, and that it holds with equality in the Gaussian case.

Proof: Here we essentially exploit statistical and information-theoretic properties of Gaussian distributions in order to transform a convex optimization problem in a conditional distribution into a convex optimization problem in a covariance matrix.

Let $\hat{w}_Q(Q)$ be the best linear MSE estimate of W from Q, and denote the covariance of the estimation error $W - \hat{w}_Q(Q)$ by

$$\Sigma_{W|Q} = \Sigma_W - \Sigma_{WQ} \Sigma_Q^{-1} \Sigma_{WQ}^{\mathrm{T}}.$$

According to the definition of the privacy-distortion function (4), we wish to minimize

$$\mathrm{I}(W; Q) = \mathrm{h}(W) - \mathrm{h}(W|Q),$$

subject to a number of constraints. Equivalently, we wish to maximize

$$\begin{aligned}\mathrm{h}(W|Q) &= \mathrm{h}(W - \hat{w}_Q(Q)|Q)\\ &\overset{\text{(a)}}{\leqslant} \mathrm{h}(W - \hat{w}_Q(Q))\\ &\overset{\text{(b)}}{\leqslant} \tfrac{1}{2} \log\left((2\pi e)^m \det \Sigma_{W|Q}\right),\end{aligned}$$

where

(a) holds with equality if and only if $W - \hat{w}_Q(Q)$ and Q are statistically independent, and

(b) follows from the fact that Gaussian r.v.'s maximize the differential entropy among all r.v.'s with a fixed covariance matrix ($\Sigma_{W|Q}$ in this case).

Proposition 3 enables us to assume without loss of generality that $\hat{X} = Q = \mathrm{E}\,[X|Q]$. But if Q is the best MSE estimate of X from Q, then the best *linear* MSE estimate $\hat{x}_Q(Q)$ of X from Q is Q itself as well. Note, however, that the reverse implication is not necessarily true. In other words, the linear constraint $\hat{x}_Q(Q) = Q$ is never more restrictive than the nonlinear one $Q = \mathrm{E}\,[X|Q]$.

To sum up,

$$\mathcal{R}(\mathcal{D}) \geqslant \mathrm{h}(W) - \max_{\substack{p_{Q|X}\\ Q=\hat{x}_Q(Q)\\ \mathrm{E}\,\|X-Q\|^2 \leqslant \mathcal{D}}} \tfrac{1}{2} \log\left((2\pi e)^m \det \Sigma_{W|Q}\right), \tag{10}$$

with equality if and only if $W - \hat{w}_Q(Q)$ and Q are statistically independent, $W - \hat{w}_Q(Q)$ is Gaussian, and $\hat{x}_Q(Q) = \mathrm{E}\,[X|Q]$. Keep in mind that all three conditions for equality are satisfied, in particular, if W, X and Q are jointly Gaussian.

We now focus our attention on the optimization problem in the right-hand side of (10), in the variable $p_{Q|X}$, after removing the superfluous constant $(2\pi e)^m$:

$$\text{maximize } \tfrac{1}{2} \log \det \Sigma_{W|Q}$$

$$\text{subject to } Q = \hat{x}_Q(Q), \text{ and } \operatorname{E} \|X - Q\|^2 \leqslant \mathcal{D},$$

and realize that the problem is completely determined by the first and second-order statistics of the r.v.'s involved. Consequently, without loss of generality, but for the exclusive purpose of solving this maximization problem, we may regard W, X and Q as jointly Gaussian, and concern ourselves solely with means and covariances. To complete the proof of the theorem, it remains to show that the maximum of this optimization problem is the same as that of the problem

$$\begin{aligned}&\text{maximize } \tfrac{1}{2} \log\det \left(\Sigma_{W|X} + \Sigma_{WX} \Sigma_X^{-1} \, \Sigma \, \Sigma_X^{-1} \Sigma_{WX}^{\mathrm{T}}\right)\\ &\text{subject to } 0 \preccurlyeq \Sigma \preccurlyeq \Sigma_X, \text{ and } \operatorname{tr} \Sigma \leqslant \mathcal{D},\end{aligned}$$

in the variable $\Sigma \in \mathbb{R}^{n \times n}$, under the assumptions that all variables involved are jointly Gaussian and, of course, that $W \leftrightarrow X \leftrightarrow Q$.

First, we verify that the constraints are equivalent, and that the variable Σ is in fact Σ_{X-Q}. From the application of the orthogonality principle of linear estimation to the constraint $Q = \hat{x}_Q(Q)$, it follows that $\Sigma_X = \Sigma_Q + \Sigma_{X-Q}$ (i.e., the observation Q and the error $X - Q$ are uncorrelated). Consequently, $0 \preccurlyeq \Sigma_{X-Q} \preccurlyeq \Sigma_X$. Conversely, since $\mu_Q = \mu_X$, any Σ_{X-Q} satisfying these matrix inequalities completely determines a $\Sigma_Q \succcurlyeq 0$ and therefore all the statistics needed to specify Q, assumed Gaussian. On the other hand, the distortion constraint may be equivalently written as $\operatorname{tr} \Sigma_{X-Q} \leqslant \mathcal{D}$.

Secondly, we check that the optimization objective is the same by expressing $\Sigma_{W|Q}$ in terms of second-order statistics of W and X. The constraint $Q = \hat{x}_Q(Q)$ implies that $\mu_Q = \mu_X$, due to the fact that optimum linear MSE estimators are unbiased, and that $\Sigma_{QX} = \Sigma_Q$, because of the orthogonality principle. Since by assumption Q and X are jointly Gaussian and $W \leftrightarrow X \leftrightarrow Q$, one may write Q as the sum of two terms. Namely, the sum of $\Sigma_{QX}\Sigma_X^{-1}(X - \mu_X) + \mu_Q$, its best linear MSE estimate from X, and the error $\mathcal{N}(0, \Sigma_{Q|X})$, which must be independent of both W and X. More precisely, defining $A = \Sigma_{Q-X}\Sigma_X^{-1}$, we have $I - A = \Sigma_Q\Sigma_X^{-1}$ and

$$\begin{aligned}Q &= \Sigma_Q\Sigma_X^{-1}(X - \mu_X) + \mu_X + \mathcal{N}(0, \Sigma_Q - \Sigma_Q\Sigma_X^{-1}\Sigma_Q)\\ &= (I - A)\, X + A\, \mathcal{N}(\mu_X, (I - A)\Sigma_X A^{-\mathrm{T}}).\end{aligned}$$

In addition, the covariance between W and Q is the same as the covariance between W and the estimate $Q = \Sigma_Q\Sigma_X^{-1}(X - \mu_X) + \mu_X$. This means that $\Sigma_{QW} = \Sigma_Q\Sigma_X^{-1}\Sigma_{XW}$, and finally,

$$\begin{aligned}\Sigma_{W|Q} &= \Sigma_W - \Sigma_{WQ}\Sigma_Q^{-1}\Sigma_{WQ}^{\mathrm{T}}\\ &= \Sigma_W - \Sigma_{WX}\Sigma_X^{-1}\Sigma_Q\Sigma_Q^{-1}\Sigma_Q\Sigma_X^{-1}\Sigma_{WX}^{\mathrm{T}}\\ &= \Sigma_W - \Sigma_{WX}\Sigma_X^{-1}(\Sigma_X - \Sigma_{X-Q})\Sigma_X^{-1}\Sigma_{WX}^{\mathrm{T}}\\ &= \Sigma_{W|X} + \Sigma_{WX}\Sigma_X^{-1}\Sigma_{X-Q}\Sigma_X^{-1}\Sigma_{WX}^{\mathrm{T}}. \qquad \blacksquare\end{aligned}$$

Appendix C
Computation of the QGLB

In this section we investigate the convex optimization problem included in the QGLB (6). Specifically, let $X, Y \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times m}$ and $t \geqslant 0$. We consider the following optimization problem in the matrix variable X:

$$\begin{aligned}&\text{maximize } \det(BB^{\mathrm{T}} + AXA^{\mathrm{T}}) &\quad (11)\\ &\text{subject to } 0 \preccurlyeq X \preccurlyeq Y, \text{ and } \operatorname{tr} X \leqslant t.\end{aligned}$$

This is in general a loose bound on the privacy-distortion function, in the sense that equality might not hold even in the Gaussian case, but it matches the QGLB for small distortion. We may solve problem (11) numerically by applying an interior-point method with log-det barrier functions [30] for the constraints $0 \preccurlyeq X \preccurlyeq Y$.

A lower bound on the QGLB, that is, a second lower bound on the privacy-distortion function, is given in Proposition 5, by means of an upper bound on problem (11). While the bound is loose in the sense that equality does not necessarily hold even in the Gaussian case, it is expressed in parametric, closed form, not requiring numerical optimization. For sufficiently low distortion, however, Proposition 6 guarantees that both bounds match. A preliminary lemma is required, which solves a convex optimization problem bearing some resemblance to the usual reverse water-filling problem of information theory, arising in the computation of the rate-distortion function for multivariate Gaussian statistics. Just as in that problem, the solution is given in parametric form.

Lemma 4: For any $r \in \mathbb{Z}^+$, let $\lambda_1 \geqslant \cdots \geqslant \lambda_r > 0$ and $y_1, \ldots, y_r, t \geqslant 0$. Consider the following optimization problem in the variables $x_1, \ldots, x_r$:

$$\begin{aligned}&\text{maximize } \prod_{i=1}^{r} (1 + \lambda_i x_i)\\ &\text{subject to } 0 \leqslant x_i \leqslant y_i \text{ for all } i, \text{ and } \sum_{i=1}^{r} x_i \leqslant t.\end{aligned}$$

The solution to the problem is given by

$$x_i = \max\left\{0, \min\left\{\frac{1}{\mu} - \frac{1}{\lambda_i}, y_i\right\}\right\} \qquad (12)$$

in the case when $\sum_{i=1} y_i > t$, and $x_i = y_i$ otherwise. In the first case, the parameter μ is chosen to satisfy $t = \sum_{i=1}^{r} x_i$. Also in this case, $\mu \geqslant \lambda_1$ implies that $x_i = 0$ for all i, and $\mu \leqslant 1/(1/\lambda_r + \max\{y_i\})$ implies that $x_i = y_i$ for all i.

Proof (Sketch): The lemma can be proved by a systematic application of the Karush-Kuhn-Tucker (KKT) conditions [30]. A more intuitive proof exploits the fact that maximizing the equivalent objective $\sum_i \ln(1 + \lambda_i x_i)$ subject to $\sum_i x_i \leqslant t$ and the rest of constraints is in fact a resource allocation problem. For instance, for all i such that $0 < x_i < y_i$, the Pareto equilibrium condition means that the marginal ratios of improvement $\frac{\mathrm{d}}{\mathrm{d}x_i} \ln(1 + \lambda_i x_i)$ must all be the same. Otherwise, minor allocation adjustments on the resources x_i could improve the objective. $\blacksquare$

The next proposition gives an upper bound on the maximum of problem (11).

Proposition 5: Let $X, Y \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times m}$ and $t \geqslant 0$. Consider the following optimization problem in the matrix variable X:

$$\begin{aligned}&\text{maximize } \det(BB^{\mathrm{T}} + AXA^{\mathrm{T}})\\&\text{subject to } 0 \preccurlyeq X \preccurlyeq Y, \text{ and } \operatorname{tr} X \leqslant t.\end{aligned}$$

Assume that B is invertible, and write the spectral decomposition of the nonnegative definite matrix $A^{\mathrm{T}}(BB^{\mathrm{T}})^{-1}A$ as $V\Lambda V^{\mathrm{T}}$, with $V = (v_1, \ldots, v_n)$ orthonormal, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$ and $\lambda_1 \geqslant \cdots \geqslant \lambda_r > 0$ (r is in fact the rank of A). Suppose further that $Y' = V^{\mathrm{T}}YV$ is diagonal. Then, the maximum in the optimization problem is

$$\det(BB^{\mathrm{T}}) \prod_{i=1}^{r} (1 + \lambda_i \xi_i), \tag{13}$$

attained by $X = \sum_{i=1}^{r} \xi_i v_i v_i^{\mathrm{T}}$, where

$$\xi_i = \max\left\{0, \min\left\{\frac{1}{\mu} - \frac{1}{\lambda_i}, y'_{ii}\right\}\right\}$$

in the case when $\sum_{i=1} y'_{ii} > t$, and $\xi_i = y'_{ii}$ otherwise. In the first case, the parameter μ is chosen to satisfy $t = \sum_{i=1}^{r} \xi_i$. Also in this case, $\mu \geqslant \lambda_1$ implies that $\xi_i = 0$ for all i, and $\mu \leqslant 1/(1/\lambda_r + \max\{y'_{ii}\})$ implies that $\xi_i = y'_{ii}$.

In the more general situation when Y' fails to be diagonal, then (13) is only an upper bound on the maximum in the optimization problem.

Proof: We proceed by gradually reducing the optimization problem of the lemma to a simpler form, finally obtaining a convex optimization problem solvable by standard techniques.

First, since B is invertible, write the optimization objective $\det(BB^{\mathrm{T}} + AXA^{\mathrm{T}})$ as

$$\det(BB^{\mathrm{T}}) \det(I + B^{-1}AX(B^{-1}A)^{\mathrm{T}}),$$

where the first factor is positive, and only the second factor depends on the variable X. Let USV^{T} be a full singular-value decomposition of $B^{-1}A$ consistent with the eigenvalue decomposition $V\Lambda V^{\mathrm{T}}$ of $(B^{-1}A)^{\mathrm{T}}B^{-1}A = A^{\mathrm{T}}(BB^{\mathrm{T}})^{-1}A$ in the statement of the lemma. In particular, this means that $S \in \mathbb{R}^{m\times n}$, with exactly r nonzero entries s_{ii} placed along the diagonal, satisfying $s_{ii} = \sqrt{\lambda_i}$, for $i = 1, \ldots, r$. Define $X' = V^{\mathrm{T}}XV$. Thus,

$$I + B^{-1}AX(B^{-1}A)^{\mathrm{T}} = I + USX'S^{\mathrm{T}}U^{\mathrm{T}} = U(I + SX'S^{\mathrm{T}})U^{\mathrm{T}}.$$

The above observations, together with the fact that U is orthonormal, lead us to conclude that maximizing the objective is equivalent to maximizing $\det(I + SX'S^{\mathrm{T}})$. It is left to formulate the optimization constraints in terms of the transformed variable X'. To this end, use the fact that V is orthonormal to write $\operatorname{tr} X' = \operatorname{tr}(V^{\mathrm{T}}XV) = \operatorname{tr} X \leqslant t$. On the other hand, since X' and Y' are defined by the same congruence transformation, the constraint $0 \preccurlyeq X \preccurlyeq Y$ is equivalent to $0 \preccurlyeq X' \preccurlyeq Y'$.

Next, we claim that the transformed optimization problem in the variable X' is in fact equivalent to the simpler maximization of $\prod_{i=1}^{r}(1 + \lambda_i x'_{ii})$ in the variables $x'_{11}, \ldots, x'_{rr}$, subject to the constraints $0 \leqslant x'_{ii} \leqslant y'_{ii}$ and $\sum_{i=1}^{r} x'_{ii} \leqslant t$, for each $i = 1, \ldots, r$. But this is precisely the optimization problem of Lemma 4, with x_i in place of x'_{ii}, renamed ξ_i in this lemma, and y_i in place of y'_{ii}.

It now remains to verify the claim under the hypothesis that Y' is diagonal. Observe first that $SX'S^{\mathrm{T}}$ is an $m \times m$ matrix which may only contain nonzero entries in the upper-left $r \times r$ block. Secondly, the constraints $\operatorname{tr} X' \leqslant t$ and $0 \preccurlyeq X' \preccurlyeq Y'$ are less restrictive than the new ones. Finally, on account of Hadamard's inequality, $\det(I + SX'S^{\mathrm{T}}) \leqslant \prod_{i=1}^{r}(1 + \lambda_i x'_{ii})$, with equality if X' is diagonal. Consequently, the optimum is precisely $X' = \operatorname{diag}(x'_{11}, \ldots, x'_{rr}, 0, \ldots, 0)$, where x'_{ii} are the solution to the last equivalent problem. To complete the proof of the lemma, observe that if Y' is not diagonal, then the solution for X' proposed may not satisfy the constraint $X' \preccurlyeq Y'$. ■

The following proposition means that for sufficiently small values of t in Proposition 5, the constraint $X \preccurlyeq Y$ will be inactive. In that case, it is clear from the proof of Proposition 5 that the solution (13) matches the QGLB.

Proposition 6: Let $X, Y \in \mathbb{R}^{n\times n}$ satisfy $X \succcurlyeq 0$ and $Y \succ 0$. Let λ^{Y}_{min} denote the minimum eigenvalue of Y. Provided that $t < \lambda^{Y}_{\text{min}}$, the constraint $\operatorname{tr} X \leqslant t$ implies that $X \prec Y$.

Proof: Since X and Y are symmetric, for any unit-norm $u \in \mathbb{R}^n$, $u^{\mathrm{T}}Yu \geqslant \lambda^{Y}_{\text{min}}$. Similarly, and since $X \succcurlyeq 0$,

$$u^{\mathrm{T}}Xu \leqslant \lambda^{X}_{\text{max}} \leqslant \sum_i \lambda^{X}_i = \operatorname{tr} X \leqslant t.$$

Hence, $u^{\mathrm{T}}(Y - X)u \geqslant \lambda^{Y}_{\text{min}} - t > 0$. ■

Acknowledgments and Disclaimers

We would like to express our gratitude to Gaetano Morrone for implementing the simulations regarding MDAV and μ-Approx in the experimental section. We would also like to thank the anonymous reviewers for numerous, invaluable comments that greatly improved the manuscript, some of them quite directly addressed.

This work was partly supported by the Spanish Government through projects CONSOLIDER INGENIO 2010 CSD2007-00004 "ARES", TSI2007-65393-C02-02 "ITACA" and TSI2007-65406-C03-01 "E-AEGIS", and by the Government of Catalonia under grants 2009 SGR 1135 and 2009 SGR 1362. The third author is partly supported by the Government of Catalonia as an ICREA Acadèmia Researcher and he is with the UNESCO Chair in Data Privacy, but his views do not necessarily reflect the position of UNESCO nor commit that organization.

References

[1] D. Rebollo-Monedero, J. Forné, and J. Domingo-Ferrer, "From t-closeness to PRAM and noise addition via information theory," in *Privacy Stat. Databases (PSD)*, ser. Lecture Notes Comput. Sci. (LNCS). Istambul, Turkey: Springer-Verlag, Sep. 2008.

[2] T. Dalenius, "Finding a needle in a haystack —or identifying anonymous census records," *J. Official Stat.*, vol. 2, no. 3, pp. 329–336, 1986.

[3] P. Samarati, "Protecting respondents' identities in microdata release," *IEEE Trans. Knowl. Data Eng.*, vol. 13, no. 6, pp. 1010–1027, 2001.

[4] P. Samarati and L. Sweeney, "Protecting privacy when disclosing information: k-Anonymity and its enforcement through generalization and suppression," SRI Int., Tech. Rep., 1998.

[5] N. Li, T. Li, and S. Venkatasubramanian, "t-Closeness: Privacy beyond k-anonymity and l-diversity," in *Proc. IEEE Int. Conf. Data Eng. (ICDE)*, Istanbul, Turkey, Apr. 2007, pp. 106–115.

[6] J. M. Gouweleeuw, P. Kooiman, L. C. R. J. Willenborg, and P.-P. de Wolf, "Post randomisation for statistical disclosure control: Theory and implementation," *J. Official Stat.*, vol. 14, no. 4, pp. 463–478, 1998.

[7] P. Kooiman, L. C. R. J. Willenborg, and J. M. Gouweleeuw, "PRAM: A method for disclosure limitation of microdata," Stat. Netherlands, Research Rep. 9705, 1997.

[8] P.-P. de Wolf, "Risk, utility and PRAM," in *Privacy Stat. Databases (PSD)*, ser. Lecture Notes Comput. Sci. (LNCS), vol. 4302. Rome, Italy: Springer-Verlag, Dec. 2006, pp. 189–204.

[9] D. Defays and P. Nanopoulos, "Panels of enterprises and confidentiality: The small aggregates method," in *Proc. Symp. Design, Anal. Longitudinal Surveys, Stat. Canada*, Ottawa, Canada, 1993, pp. 195–204.

[10] J. Domingo-Ferrer and J. M. Mateo-Sanz, "Practical data-oriented microaggregation for statistical disclosure control," *IEEE Trans. Knowl. Data Eng.*, vol. 14, no. 1, pp. 189–201, 2002.

[11] J. Domingo-Ferrer and V. Torra, "Ordinal, continuous and heterogenerous k-anonymity through microaggregation," *Data Min., Knowl. Disc.*, vol. 11, no. 2, pp. 195–212, 2005.

[12] J. Domingo-Ferrer, F. Sebé, and A. Solanas, "A polynomial-time approximation to optimal multivariate microaggregation," *Comput., Math. with Appl.*, vol. 55, no. 4, pp. 714–732, Feb. 2008.

[13] T. M. Truta and B. Vinay, "Privacy protection: p-sensitive k-anonymity property," in *Proc. Int. Workshop Privacy Data Manage. (PDM)*, Atlanta, GA, 2006, p. 94.

[14] X. Sun, H. Wang, J. Li, and T. M. Truta, "Enhanced p-sensitive k-anonymity models for privacy preserving data publishing," *Trans. Data Privacy*, vol. 1, no. 2, pp. 53–66, 2008.

[15] A. Machanavajjhala, J. Gehrke, D. Kiefer, and M. Venkitasubramanian, "l-Diversity: Privacy beyond k-anonymity," in *Proc. IEEE Int. Conf. Data Eng. (ICDE)*, Atlanta, GA, Apr. 2006, p. 24.

[16] J. Domingo-Ferrer and V. Torra, "A critique of k-anonymity and some of its enhancements," in *Proc. Workshop Privacy, Security, Artif. Intell. (PSAI)*, Barcelona, Spain, 2008, pp. 990–993.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[18] J. Brickell and V. Shmatikov, "The cost of privacy: Destruction of data-mining utility in anonymized data publishing," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc., Data Min. (KDD)*, Las Vegas, NV, Aug. 2008.

[19] A. Evfimievski, J. Gehrke, and R. Srikant, "Limiting privacy breaches in privacy preserving data mining," in *Proc. ACM Symp. Prin. Database Syst. (PODS)*, San Diego, CA, 2003, pp. 211–222.

[20] C. E. Shannon, "Communication theory of secrecy systems," Bell Syst., Tech. J., 1949.

[21] A. Wyner, "The wiretap channel," Bell Syst., Tech. J. 54, 1975.

[22] P. M. Woodward, "Theory of radar information," in *Proc. London Symp. Inform. Theory, Ministry of Supply*, London, UK, 1950, pp. 108–113.

[23] D. V. Lindley, "On a measure of the information provided by an experiment," *Annals Math. Stat.*, vol. 27, no. 4, pp. 986–1005, 1956.

[24] A. G. de Waal and L. C. R. J. Willenborg, "Information loss through global recoding and local suppression," *Netherlands Official Stat.*, vol. 14, pp. 17–20, 1999.

[25] L. Willenborg and T. DeWaal, *Elements of Statistical Disclosure Control*. New York: Springer-Verlag, 2001.

[26] E. Bertino and M. L. Damiani, "Foreword for the special issue of selected papers from the 1$^{\text{st}}$ ACM SIGSPATIAL Workshop on Security and Privacy in GIS and LBS," *Trans. Data Privacy*, vol. 2, no. 1, pp. 1–2, 2009.

[27] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," in *IRE Nat. Conv. Rec.*, vol. 7 Part 4, 1959, pp. 142–163.

[28] P. Venkitasubramaniam and L. Tong, "A rate-distortion approach to anonymous networking," in *Proc. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, Sep. 2007.

[29] P. Venkitasubramaniam, T. He, and L. Tong, "Anonymous networking amidst eavesdroppers," *IEEE Trans. Inform. Theory, Special Issue Inform.-Theor. Security*, vol. 54, no. 6, pp. 2770–2784, Jun. 2008.

[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[31] J. Domingo-Ferrer, A. Martínez-Ballesté, J. M. Mateo-Sanz, and F. Sebé, "Efficient multivariate data-oriented microaggregation," *VLDB J.*, vol. 15, no. 4, pp. 355–369, 2006.

[32] A. Hundepool, R. Ramaswamy, P.-P. DeWolf, L. Franconi, R. Brand, and J. Domingo-Ferrer, "μ-ARGUS version 4.1 software and user's manual," Voorburg, Netherlands, 2007. [Online]. Available: http://neon.vb.cbs.nl/casc

[33] M. Templ, "Statistical disclosure control for microdata using the R-package sdcMicro," *Trans. Data Privacy*, vol. 1, no. 2, pp. 67–85, 2008. [Online]. Available: http://cran.r-project.org/web/packages/sdcMicro

David Rebollo-Monedero received the M.S. and Ph.D. degrees in electrical engineering from Stanford University, in California, USA, in 2003 and 2007, respectively. His doctoral research at Stanford focused on data compression, more specifically, quantization and transforms for distributed source coding. Previously, he was an information technology consultant for PricewaterhouseCoopers, in Barcelona, Spain, from 1997 to 2000, and was involved in the Retevisión startup venture. He is currently a Postdoctoral Researcher with the Information Security Group in the Department of Telematics of the Technical University of Catalonia (UPC), also in Barcelona.

Jordi Forné received the M.S. degree in telecommunications engineering from the Technical University of Catalonia (UPC) in 1992 and the Ph.D. degree in 1997. In 1991, he joined the Cryptography and Network Security Group at the Department of Applied Mathematics and Telematics. Currently, he is with the Information Security Group at the Department of Telematics Engineering of the UPC. His research interests span a number of subfields within information security and privacy, including network security, electronic commerce and public key infrastructures. Currently, he works as an Associate Professor at the Telecommunications Engineering School at UPC.

Josep Domingo-Ferrer is a Full Professor of Computer Science and an ICREA-Acadèmia Researcher at Universitat Rovira i Virgili, Tarragona, Catalonia, where he holds the UNESCO Chair in Data Privacy. He received with honors his M.S. and Ph.D. degrees in Computer Science from the Universitat Autònoma de Barcelona in 1988 and 1991, respectively (Outstanding Graduation Award). He also holds a M.S. in Mathematics. His fields of activity are data privacy, data security and cryptographic protocols. He has received three research awards and four entrepreneurship awards, among which the ICREA Acadèmia Research Prize from the Government of Catalonia. He has authored 3 patents and over 220 publications, one of which became an ISI highly-cited paper in early 2005. He has been the coordinator of EU FP5 project CO-ORTHOGONAL and of several Spanish funded and U.S. funded research projects. He currently coordinates the CONSOLIDER "ARES" team on security and privacy, one of Spain's 34 strongest research teams. He has chaired or cochaired 9 international conferences and has served in the program committee of over 70 conferences on privacy and security. He is a co-Editor-in-Chief of "Transactions on Data Privacy" and an Associate Editor of three international journals. In 2004, he was a Visiting Fellow at Princeton University.