Differential Privacy Trough Knowledge Refinement

Jordi Soria-Comas and Josep Domingo-Ferrer
UNESCO Chair in Data Privacy
Dept. of Computer Engineering and Maths
Universitat Rovira i Virgili
Av. Països Catalans 26, E-43007 Tarragona, Catalonia
E-mail {jordi.soria,josep.domingo}@urv.cat

***Abstract*—We introduce a novel mechanism to attain differential privacy. Contrary to the common mechanism based on the addition of a noise whose magnitude is proportional to the sensitivity of the query function, our proposal is based on the refinement of the user's prior knowledge about the response. We show that our mechanism has several advantages over noise addition: it does not require complex computations, and thus it can be easily automated; it lets the user exploit her prior knowledge about the response to achieve better data quality; and it is independent of the sensitivity of the query function (although this can be a disadvantage if the sensitivity is small). We also show some compounding properties of our mechanism for the case of multiple queries.**

***Index Terms*—Differential privacy, knowledge refinement, statistical databases.**

I. Introduction

Differential privacy [1], [2] is a statistical disclosure control (SDC) methodology based on output perturbation. The disclosure risk limitation offered by differential privacy is based on the limitation of the effect that any single individual has on a query response. If the influence of any single individual on the query response is small, publishing that response involves only a small disclosure risk for any individual.

Any mechanism used to achieve differential privacy may be seen as the application of a perturbation to the real value of the query response. The original proposal [1], [2] to attain differential privacy masks the query response by adding a Laplace distributed noise whose magnitude is proportional to the global sensitivity of the query function. If the sensitivity of the query function is not constant across data sets, the magnitude of the noise must be selected to offer the required level of disclosure limitation to the data set with the greatest sensitivity, thereby overprotecting the rest of data sets. The approach in [3] tries to avoid this problem by adjusting for each data set the magnitude of the noise to the local sensitivity of the query function. Other mechanisms proposed to attain differential privacy include the exponential mechanism [4], that introduces the concept of response utility, and some mechanisms designed for specific types of queries, such as [5], [6].

All the methods presented are at some point concerned with the variability (a.k.a. sensitivity) of the query function between neighbor data sets. This kind of mechanisms present two main problems: (i) deciding what amount of noise to add may require complex computations (such as computing the global sensitivity [1] or the smooth sensitivity [3] of the query function) and thus noise addition may be difficult to automate, and (ii) in some cases, the amount of noise that needs to be added is so large that the output may bear little resemblance to the real query response, and thus the response may be misleading.

A. Contribution of this paper

We introduce a mechanism to achieve differential privacy that works by refining the prior knowledge/beliefs of the database user as much as possible, given the constraints set by differential privacy. Our mechanism depends only on the prior knowledge and on the level of protection that we want to achieve. It is completely independent from the actual query function, and thus no complex sensitivity computations are required.

This mechanism avoids problem (i) above, as no complex computations are required. Regarding problem (ii), the mechanism guarantees that the response provides increased utility over the prior knowledge that the user had.

Section II describes the proposed knowledge refinement mechanism, both for continuous and discrete responses, and compares it with Laplace noise addition. Section III generalizes the approach for compounded queries. Section IV contains a discussion and Section V summarizes conclusions.

II. Refining prior knowledge to achieve differential privacy

The definition of differential privacy states that the probability for the response to belong to some set must be similar regardless of whether any specific individual is included or not in the data set.

Definition 1 (ε-differential privacy, [2])**.** A randomized function κ gives ε-differential privacy if, for all data sets D_1, D_2 such that one can be obtained from the other by adding or removing a single record, and all $S \subset Range(\kappa)$

$$P(\kappa(D_1) \in S) \leq e^{\varepsilon} \times P(\kappa(D_2) \in S) \tag{1}$$

A common approach to satisfy the requirements of Definition 1 is noise addition: first, the real value of the query response is computed and, then, a random noise is added to mask it. A Laplace distribution with zero mean and a scale

parameter that depends on the variability of the query function is commonly used for noise addition.

Our proposal is not based on masking the true value of the response by adding some noise, but on modifying the prior knowledge of the database user on the response. When a query is submitted to the database, the user submits at the same time her knowledge/beliefs about the response. We think of this prior knowledge as the probability distribution that the user expects for the response. For example, in case the user has absolutely no idea about the possible result for a query f, the probability distribution to use is the uniform distribution over the range of f (assuming that this range is bounded). The access mechanism modifies this prior knowledge to fit the real value of the response as much as possible given the constraints imposed by differential privacy.

Definition 2 (Prior knowledge)**.** Given a query function f, the *prior knowledge* about the response $f(D)$ is the probability distribution P_f, defined over $Range(f)$, that the user expects for the response to f.

The more concentrated the probability mass of P_f around the real value of the response to f, the more accurate is the user's prior knowledge. In general, as the user knows the query f, and the set of possible databases $\mathcal{D}$, we may expect some knowledge about the response $f(D)$. The better the knowledge the user has on the actual database D, the more accurate is the prior knowledge the user can provide to the response mechanism.

If the query function f has multiple components (dimension $n > 1$), the joint probability distribution must be provided. If the components of f are independent, specifying the marginal distribution for each component is enough to compute the joint distribution. This will also be the case if the components are not independent but the user has no knowledge about the relationship among them.

The access mechanism works as follows:

1) Receive the query f and the prior knowledge P_f from the database user.
2) Compute the actual value of the query response, $f(D)$.
3) Modify P_f to adjust it to $f(D)$ as much as possible, given the constraints imposed by differential privacy.
4) Randomly sample the distribution resulting from the previous step, and return the sampled value as the response to f evaluated at D.

The critical step is the adjustment of the prior knowledge to the real query response. To perform this adjustment, we distinguish two types of queries: statistical queries and individual queries. We call statistical queries those whose outcome depends on multiple individuals, while individual queries are those that depend on a single individual. It will be shown below that a finer adjustment of the prior knowledge is feasible for individual queries. We start by focusing on statistical queries, but, before formally specifying the response mechanism, we provide an example to illustrate what we intend to do.

Example 1. Assume a query function f that is known to return a value within the interval $[0, 1]$. Assume also that the database user has no further knowledge about the query response, *i.e.* her prior knowledge is $\mathcal{U}[0, 1]$, the uniform distribution over $[0, 1]$.

To refine the prior knowledge, we modify its density by applying a multiplicative factor $\alpha_u \geq 1$ to the points near $f(D)$, and a factor $\alpha_d \leq 1$ to the points farther from $f(D)$. In this way, the probability of obtaining as response a value near the actual response $f(D)$ is increased, with respect to the prior knowledge, while the probability of obtaining a distant value is decreased. Figure 1 shows the probability distribution resulting from applying the procedure described above for a couple of neighbor datasets D and D'.

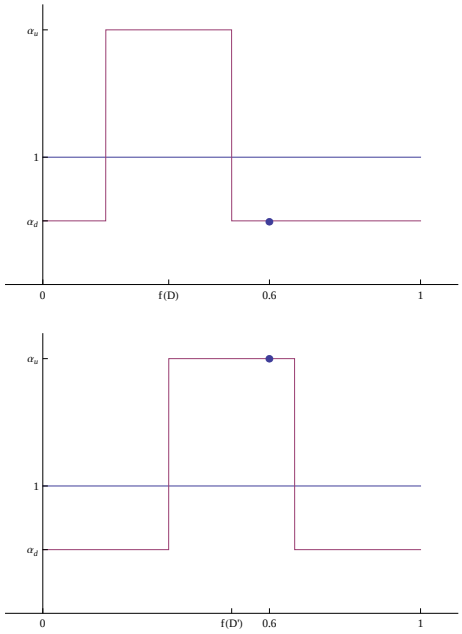

Fig. 1. Distributions for the response to $f(D)$ (left) and to $f(D')$ (right)

To obtain ε-differential privacy, the density at a given point for the response to $f(D)$ must be a factor within the interval $[e^{-\varepsilon}, e^{\varepsilon}]$ of the density at the same point for the response to $f(D')$. Check, for example, the point 0.6 in Figure 1: on the left-hand side distribution, the point is far from the real response and thus a factor α_d is applied; on the right-hand side distribution, the point is near the real response and the factor applied is α_u. For the ε-differential privacy condition to hold, it must be $\alpha_u/\alpha_d \leq e^{\varepsilon}$. We can also think in the reverse way: given a couple of constants $\alpha_u \geq 1$ and $\alpha_d \leq 1$, the level of differential privacy achieved by this response mechanism is $\varepsilon = \ln(\alpha_u/\alpha_d)$.

Note that, to obtain a valid density function from the above modification, the set of points over which α_u and α_d is applied must be selected in such a way that the total probability mass of the resulting distribution equals 1. If we denote $\mathcal{U}_u$ as the set over which we apply the factor α_u, for the total probability mass of the adjusted distribution to be 1, we must

have $\alpha_u P_f(\mathcal{U}_u) + \alpha_d(1 - P_f(\mathcal{U}_u)) = 1$. If the prior knowledge is an absolutely continuous distribution, as in Example 1, for any pair of values $\alpha_u \geq 1$, $\alpha_d \leq 1$ it is possible to select a set $\mathcal{U}_u$ in such a way that $\alpha_u P_f(\mathcal{U}_u) + \alpha_d(1 - P_f(\mathcal{U}_u)) = 1$ is satisfied. The reason is that we can select the set $\mathcal{U}_u$ to have any probability mass between 0 and 1. If the prior knowledge distribution is not absolutely continuous, it may not be possible to define a set $\mathcal{U}_u$ with the required probability mass for the given values α_u and α_d, and therefore we may need to adjust the values of α_u and α_d. If we call $\tilde{\alpha}_u$ and $\tilde{\alpha}_d$ the adjusted values, the level of differential privacy we achieve is $\ln(\tilde{\alpha}_u/\tilde{\alpha}_d)$. In Example 2 we will see how the values α_u and α_d must be adjusted for a discrete distribution.

The following proposition formalizes the ideas exposed in the previous example.

Proposition 1. *Let $f{:}\mathcal{D} \to \mathbb{R}^n$ be a query function, and let P_f be the prior knowledge distribution for f. Let $\alpha_u \geq 1$ and $\alpha_d \leq 1$ be such that $\alpha_u = e^{\varepsilon}\alpha_d$. Let $\mathcal{U}_u$ be an environment of $f(D)$ satisfying $\alpha_u P_f(\mathcal{U}_u) + \alpha_d(1 - P_f(\mathcal{U}_u)) = 1$. The response mechanism that returns a value randomly sampled from the distribution obtained by modifying P_f through multiplication of the probability mass of the points in $\mathcal{U}_u$ by α_u, and multiplication of the probability mass of the points outside $\mathcal{U}_u$ by α_d, satisfies ε-differential privacy.*

When the query f returns a value related to a single individual, the mechanism in Proposition 1 can be improved. In such a case, there are only two possibilities for the response: (i) if the individual we are asking for is not in the database, the distribution of the response equals the prior knowledge distribution, and (ii) if the individual is the database, the distribution for the response will be the result from the refinement of the prior knowledge. To satisfy ε-differential privacy, we only need to guarantee that the distribution resulting from (i) and (ii) does satisfy the limitation on the knowledge gain imposed by differential privacy. In other words, the output distribution need only be compared to the prior knowledge. The conditions that must hold are $1 \leq \alpha_u \leq e^{\varepsilon}$ and $e^{-\varepsilon} \leq \alpha_d \leq 1$.

Note that, by choosing $\alpha_u = e^{\varepsilon}$ and $\alpha_d = e^{-\varepsilon}$, the level of differential privacy that we can guarantee for a statistical query function (depending on multiple individuals) is 2ε, while for an individual query (whose outcome depends on a single individual), we double the guarantee to ε.

Proposition 2. *Let $f{:}\mathcal{D} \to \mathbb{R}^n$ be an individual query in the above sense, and let P_f be the prior knowledge distribution for f. Let $\alpha_u = e^{\varepsilon}$ and $\alpha_d = e^{-\varepsilon}$. Let $\mathcal{U}_u$ be an environment of $f(D)$ satisfying $\alpha_u P_f(\mathcal{U}_u) + \alpha_d(1 - P_f(\mathcal{U}_u)) = 1$. The response mechanism that returns a value randomly sampled from the distribution obtained by modifying P_f through multiplication of the probability mass of the points in $\mathcal{U}_u$ by α_u, and multiplication of the probability mass of the points outside $\mathcal{U}_u$ by α_u, satisfies ε-differential privacy.*

Same as with Proposition 1, for the case of a non absolutely continuous prior knowledge distribution, we may need to adjust the values of α_u and α_d in such a way that $\alpha_u P_f(\mathcal{U}_u) + \alpha_d(1 - P_f(\mathcal{U}_u)) = 1$ holds.

To give some idea of the kind of results that we could expect from the proposed mechanism, we compare it against the typical Laplace noise addition mechanism for the case of a simple individual query that returns the value of a Boolean attribute.

Example 2 (Data quality for a Boolean attribute)**.** Consider a simple database D with two attributes: an identifier ID, and a Boolean attribute B that may take the values 0 and 1. We assume that B is very sensitive and that, to limit the disclosure risk, access to the database must be mediated by a query-response mechanism satisfying differential privacy, with $\varepsilon = 1$. Let $f : \mathcal{D} \to \{0, 1\}$ be a query that asks the value in the attribute B for a specific individual.

To achieve differential privacy via Laplace noise addition, we must first compute the sensitivity of function f. Assuming that f returns $1/2$ if the individual is not in the database, the L_1-sensitivity of f is $1/2$. Therefore, to achieve differential privacy for $\varepsilon = 1$, we must add Laplace distribution $L(0, 1/2)$ to the true value of the query response. Figure 2 shows the distribution of the responses for both possible values of B, 0 and 1.

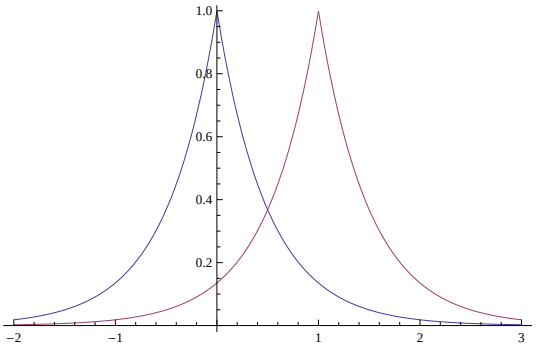

Fig. 2. Response distributions with Laplace noise addition

Assuming that the user is only interested in a 0-1 response, any value below $1/2$ is taken as 0, and any value above $1/2$ as 1. The distribution for the response thus obtained is:

$$K_f(D) = \begin{cases} 0 & f(D) + L(0, 1/2) < 0.5 \\ 1 & otherwise \end{cases}$$

If $f(D)$ equals 0, $K_f(D)$ follows a Bernoulli distribution with parameter 0.184. If $f(D)$ equals 1, the distribution of $K_f(D)$ is a Bernoulli with parameter 0.816. Note that this is completely independent from the true distribution of attribute B, and from any prior knowledge that the user might have on it. Hence, differential privacy via Laplace noise addition does not allow the user to exploit prior knowledge.

Let us assume that attribute B is 1 only with probability 0.01. For a user with this information, using the response obtained from the differential privacy mechanism is actually misleading, as the result will be 1 with probability

$$P(K_f(D) = 1 | f(D) = 0) P_f(0) +$$

$$+P(K_f(D)=1|f(D)=1)P_f(1)$$

$$=0.184\cdot 0.99+0.816\cdot 0.01=0.19$$

We could increase the parameter ε to get a more accurate response. However, by doing so we would be reducing the privacy guarantees.

Now, we turn to the refinement mechanism and, same as before, we assume that the user knows that B equals 1 with probability 0.01. Take $\alpha_u = e^{\varepsilon} = e$ and $\alpha_d = e^{-\varepsilon} = e^{-1}$. Now

$$P(K_f(D)=1|f(D)=0)=P_f(1)\cdot\alpha_d=0.003678$$

$$P(K_f(D)=0|f(D)=0)=1-0.003678=0.9963222$$

$$P(K_f(D)=1|f(D)=1)=P_f(1)\cdot\alpha_u=0.027182$$

$$P(K_f(D)=0|f(D)=1)=1-0.027182=0.972817$$

Note that, as this is not an absolutely continuous distribution, we had to do some adjustment to have a total probability mass equal to one: instead of adjusting α_u and α_d, we directly adjusted $P(K_f(D)=0|f(D)=0)$ and $P(K_f(D)=0|f(D)=1)$. Figure 3 depicts the distribution of the response for both possible values of attribute B and for the prior knowledge.

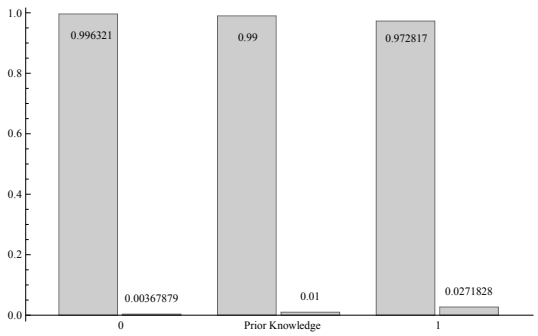

Fig. 3. Response distributions with prior knowledge refinement

Now, the probability of obtaining a response 1 is

$$P(K_f(D)=1)=$$

$$=P(K_f(D)=1|f(D)=0)P_f(0)+$$

$$+P(K_f(D)=1|f(D)=1)P_f(1)$$

$$=0.003678\cdot 0.99+0.027182\cdot 0.01=0.003912$$

As 0.003912 is much closer to 0.01 than 0.19, we conclude that, despite both mechanisms providing the same level of privacy, the output distribution is much closer to the actual distribution of the attribute when using the mechanism based on knowledge refinement. Therefore, the knowledge refinement mechanism is more suitable for a micro-data release than the Laplace noise addition.

III. Compounded queries

The above knowledge refinement mechanism can be compounded in the sense that if we obtain an ε_1-differentially private response for a query f_1 and an ε_2-differentially private response for a query f_2, we have an $(\varepsilon_1+\varepsilon_2)$-differentially private response for the query (f_1, f_2). This is in fact a property of ε-differential privacy, hence a proof for our specific mechanism is not required (see [4]). More generally, n ε-differentially private responses to n corresponding queries can be viewed as an $n\varepsilon$-differentially private response to the compounded query.

The above result on compounded queries can be improved when each of the queries refers to a disjoint set of individuals. For the noise addition mechanism, it easy to see that, when performing queries $f_1,\ldots,f_n$ that refer each to a disjoint set of individuals, the global sensitivity equals the maximum of the sensitivities of the individual queries [1]. The reason is that, by adding or removing a single individual from the data set, only one of the queries is affected. This is a good property, as it guarantees $\max\{\varepsilon_i\}$-differential privacy instead of $\sum\varepsilon_i$-differential privacy. Our goal is to show that this property can also be achieved for our proposal. We start with an example.

Example 3. Let D be a database with two attributes: an identifier ID, and a Boolean attribute B. Let f_1 and f_2 be queries that return the value of B for individuals 1 and 2, respectively. Let the prior knowledge for both queries be the independent uniform distributions over the set $\{0,1\}$, which assign a prior probability 0.5 to each of the possible outcomes for each query. To respond to f_1 in an ε-differentially private way with $\varepsilon=1$, we select factors $\alpha_u = e^{\varepsilon}$ and $\alpha_d = e^{-\varepsilon}$ that modify the prior knowledge. The same factors are selected for f_2. Now we want to check whether the combination of responses to f_1 and f_2 is still ε-differentially private.

For the sake of simplicity, we assume that both individuals are in D, and that $f_1(D)=0$ and $f_2(D)=0$. For the rest of cases we would proceed in a similar way. Figure 4 shows the prior knowledge and the distribution of the ε-differentially private response to query functions f_1 and f_2. Indeed, setting $\alpha_d = e^{-1}$ and adjusting the probability mass to 1 instead of setting $\alpha_u = e^1$, we have

$$P(K_{f_1}(D)=1|f_1(D)=0)=P(K_{f_2}(D)=1|f_2(D)=0)=$$

$$=0.5\alpha_d=0.5e^{-1}=0.183939$$

$$P(K_{f_1}(D)=0|f_1(D)=0)=P(K_{f_2}(D)=0|f_2(D)=0)=$$

$$=1-0.5\alpha_d=0.816061$$

Table I shows the joint distribution for the output of (f_1, f_2), which is obtained by multiplying the output distributions for f_1 and f_2.

For ε-differential privacy to hold for the compounded query $f=(f_1,f_2)$, the ratio of the response distribution at D and the response distribution at any D' that results from from D by adding or removing a single individual must be within the range $[e^{-\varepsilon}, e^{\varepsilon}]$. As f_1 and f_2 are related to individuals 1 and 2,

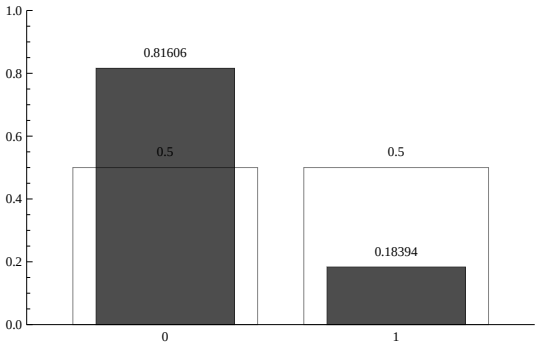

Fig. 4. Prior knowledge about attribute B and distribution of the ε-differentially private response to query functions f_1 and f_2, assuming that the actual value for attribute B is 0

TABLE I
DISTRIBUTION OF THE DIFFERENTIALLY PRIVATE RESPONSE TO THE COMPOUNDED QUERY (f_1, f_2) WHEN THE TRUE VALUES ARE $f_1(D) = f_2(D) = 0$

K_{f_2} \ K_{f_1}	0	1
	$1-0.5\alpha_d$	$0.5\alpha_d$
0 $\quad 1-0.5\alpha_d$	$(1-0.5\alpha_d)^2$	$(1-0.5\alpha_d)0.5\alpha_d$
1 $\quad 0.5\alpha_d$	$(1-0.5\alpha_d)0.5\alpha_d$	$0.25\alpha_d^2$

any modification to D that does not affect the records for those individuals leaves the distribution of responses unchanged. As we are assuming that individuals 1 and 2 are in D, the only modifications to consider are the removal of one of these individuals. Tables II show the distributions of responses when individual 1 or 2 are removed. We use K_f to denote the distribution of the response to query f. It can be seen that the respective ratios between the distribution in Table I and the ones in Table II are within $[e^{-\varepsilon}, e^{\varepsilon}] = [e^{-1}, e]$; specifically, the ratios take only two values, $\alpha_d = e^{-1}$ and $2 - \alpha_d = 2 - e^{-1}$.

TABLE II
DISTRIBUTION OF THE RESPONSE TO QUERY $f = (f_1, f_2)$ WHEN EITHER INDIVIDUAL 1 IS MISSING (LEFT) OR INDIVIDUAL 2 IS MISSING (RIGHT), AND WHEN THE ATTRIBUTE VALUE FOR THE NON-MISSING INDIVIDUAL IS 0

K_{f_2} \ K_{f_1}	0	1
	0.5	0.5
0 $\quad 1-0.5\alpha_d$	$0.5(1-0.5\alpha_d)$	$0.5(1-0.5\alpha_d)$
1 $\quad 0.5\alpha_d$	$0.25\alpha_d$	$0.25\alpha_d$

K_{f_2} \ K_{f_1}	0	1
	$1-0.5\alpha_d$	$0.5\alpha_d$
0 $\quad 0.5$	$0.5(1-0.5\alpha_d)$	$0.25\alpha_d$
1 $\quad 0.5$	$0.5(1-0.5\alpha_d)$	$0.25\alpha_d$

Proposition 3. *Let D be a data set, and let $(f_1, \ldots, f_n)$ be a set of query functions referring to disjoint sets of individuals. Let K_{f_i} be a random variable that provides ε_i-differential privacy for f_i, and assume that K_{f_i} is independent from K_{f_j} for any $i \neq j$. Then $(K_{f_1}, \ldots, K_{f_n})$ provides $\max\{\varepsilon_i\}$-differential privacy for $(f_1, \ldots, f_n)$.*

Proof: Let D' be a data set obtained from D by adding or removing a single user. We want to check that the following inequalities hold for any subset S of the range of $(K_{f_1}, \ldots, K_{f_n})$:

$$e^{-\max\{\varepsilon_i\}} \leq \frac{P((K_{f_1}(D), \ldots, K_{f_n}(D)) \in S)}{P((K_{f_1}(D'), \ldots, K_{f_n}(D')) \in S)} \leq e^{\max\{\varepsilon_i\}}$$

It is easy to show that the above inequality holds for the case of S being the Cartesian product of sets S_i, with S_i a subset of the range of $K_{f_i}(D)$, or when the probability distribution of $(K_{f_1}, \ldots, K_{f_n})$ is absolutely continuous. For a general set S and a non absolutely continuous distribution, the inequalities still hold. However, such a general proof requires the use of some concepts of measure theory and, for space reasons, we omit such details here. We will show that the inequalities hold for the case of $S = S_1 \times \ldots \times S_n$.

The probabilities $P((K_{f_1}(D), \ldots, K_{f_n}(D)) \in S)$ and $P((K_{f_1}(D'), \ldots, K_{f_n}(D')) \in S)$ can be written as the product of probabilities $\prod P(K_{f_i}(D) \in S_i)$ and $\prod P(K_{f_i}(D') \in S_i)$, respectively. By adding or removing a single individual, only one of the queries is affected. Say the affected query is f_j for some $j \in \{1, \ldots, n\}$. By removing the factors that are both in the numerator and the denominator, the inequalities that we need to check become:

$$e^{-\max\{\varepsilon_i\}} \leq \frac{P(K_{f_j}(D) \in S_j)}{P(K_{f_j}(D') \in S_j)} \leq e^{\max\{\varepsilon_i\}}$$

which hold because K_{f_j} satistifies ε_j-differential privacy, and $\varepsilon_j \leq \max\{\varepsilon_i\}$. ■

IV. DISCUSSION

In the examples in the previous sections, we have highlighted that the mechanism based on prior knowledge refinement lets the database user exploit any prior knowledge about the query response. This yields responses which have better quality than with noise addition while still being differentially private.

Other advantages of prior knowledge refinement are:

- Simplicity. Mechanisms such as Laplace noise addition are based on the addition of a random noise whose magnitude depends on the variability of the query function across neighbor data sets, a.k.a. sensitivity. To calibrate the random noise, the sensitivity of the function must be computed, which may be quite complex. The mechanism based on the refinement of the prior knowledge only depends on the prior knowledge (it is independent from the sensitivity of the query function), and thus it is easier to implement, especially in a non-supervised environment.
- Generality. As said above, Laplace noise addition requires computing the sensitivity of the query function, and this can only be done if the query function takes values in a metric space. This introduces some complexities when the function returns categorical information. The mechanism based on prior knowledge refinement does not impose any requirement on the query function, and thus it can

be applied without extra overhead to functions returning categorical information.

Despite the advantages listed above, there are some situations for which the proposed mechanism is not appropiate. If the range of values that the function may return is large compared to the variability between neighbor data sets, and the database user does not have precise knowledge of the response, then a method based on noise addition produces better data quality. This may be the case of statistical queries where the user has no prior knowledge of the result. However, when querying a specific individual, the proposed method results in much greater response quality.

V. Conclusions

We have introduced a novel mechanism to attain differential privacy. This mechanism is based on refining the prior knowledge that the user may have about the query response. This refinement is performed taking into account the constraints imposed by differential privacy.

The mechanism presents several advantages over the usual noise addition mechanism. It is easier to implement, especially in a non-supervised environment, as it does not require potentially complex computations (such as determining the sensitivity of the query function). The fact that it lets users exploit their prior knowledge may lead to a level of data quality not reachable by mechanisms independent of the user knowledge. For example, we showed in Example 2 that the distribution of the response was closer to the actual distribution when using the proposed mechanism. For query functions with large sensitivity, the amount of noise added by noise addition mechanisms, such as [1], may render the response useless. In contrast, the data quality that results from our proposal is independent from the sensitivity of the query function (yet this has the drawback that, for small sensitivities, our approach may be inferior to noise addition).

We have also analyzed the behavior of our approach for compounded, *i.e.* multiple, queries. A generic property of differential privacy guarantees that, if an ε_i-differentially private response is provided to a query f_i, for $i = 1$ to n, a $\sum \varepsilon_i$-differentially private response is provided to the compounded query $(f_1, \dots f_n)$. We have seen that this can be improved if each query f_i refers to a disjoint set of individuals. In this case, we achieve $\max\{\varepsilon_i\}$-differential privacy, instead of $\sum \varepsilon_i$-differential privacy.

Acknowledgments and disclaimer

This work was partly supported by the Government of Catalonia under grant 2009 SGR 1135, by the Spanish Government through projects TSI2007-65406-C03-01 "E-AEGIS", TIN2011-27076-C03-01 "CO-PRIVACY" and CONSOLIDER INGENIO 2010 CSD2007-00004 "ARES", and by the European Comission under FP7 projects "DwB" and "Inter-Trust". The second author is partially supported as an ICREA Acadèmia researcher by the Government of Catalonia. The authors are with the UNESCO Chair in Data Privacy, but they are solely responsible for the views expressed in this paper, which do not necessarily reflect the position of UNESCO nor commit that organization.

References

[1] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the 3rd Theory of Cryptography Conference*, LNCS 3876, pages 265–284. Springer, 2006.

[2] C. Dwork. Differential privacy. In *Automata, Languages and Programming*, LNCS 4052, pages 1–12. Springer, 2006.

[3] K. Nissim, S. Raskhodnikova, and A. Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the 39th Annual ACM Symposium on the Theory of Computing - STOC '07*, pages 75–84. ACM, 2007.

[4] F. McSherry and K. Talwar. Mechanism design via differential privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 94–103. IEEE, 2007.

[5] B. Barak, S. Kale, K. Chaudhuri, F. Mcsherry, C. Dwork and K. Talwar Privacy, accuracy, and consistency too: a holistic solution to contingency table release. In *Proceedings of the 26th Symposium on Principles of Database Systems - PODS '07*, pages 273–282. 2007.

[6] X. Xiao, G. Wang and J. Gehrke. Differential privacy via wavelet transforms. In *Proceedings of the 26th International Conference on Data Engineering - ICDE '10*, pages 225–236, 2010.