Impossibility of Differentially Private Universally Optimal Mechanisms

Hai Brenner
dept. of Mathematics
Ben-Gurion University
Beersheba, Israel
haib@bgu.ac.il

Kobbi Nissim
dept. of Computer Science
Ben-Gurion University
Beersheba, Israel
kobbi@cs.bgu.ac.il

***Abstract*—The notion of *a universally utility-maximizing privacy mechanism* was recently introduced by Ghosh, Roughgarden, and Sundararajan [STOC 2009]. These are mechanisms that guarantee optimal utility to a large class of information consumers, *simultaneously*, while preserving *Differential Privacy* [Dwork, McSherry, Nissim, and Smith, TCC 2006]. Ghosh, Roughgarden and Sundararajan have demonstrated, quite surprisingly, a case where such a universally-optimal differentially-private mechanisms exists, when the information consumers are Bayesian. This result was recently extended by Gupte and Sundararajan [PODS 2010] to risk-averse consumers.**

Both positive results deal with mechanisms (approximately) computing a *single count query* (i.e., the number of individuals satisfying a specific property in a given population), and the starting point of our work is a trial at extending these results to similar settings, such as sum queries with non-binary individual values, histograms, and two (or more) count queries. We show, however, that universally-optimal mechanisms do not exist for all these queries, both for Bayesian and risk-averse consumers.

For the Bayesian case, we go further, and give a characterization of those functions that admit universally-optimal mechanisms, showing that a universally-optimal mechanism exists, essentially, only for a (single) count query. At the heart of our proof is a representation of a query function f by its *privacy constraint graph* G_f whose edges correspond to values resulting by applying f to neighboring databases.

***Keywords*-differential privacy; universally optimal mechanisms; utility; geometric mechanism;**

I. Introduction

Differential Privacy [6] is a rigorous notion of privacy that allows learning global ('holistic') information about a collection of individuals while preserving each individual's information private. The literature of differential privacy is now rich in techniques for constructing differentially privacy mechanisms, including some generic techniques such as the addition of Laplace noise with magnitude calibrated to global sensitivity [6], addition of instance based noise calibrated to smooth sensitivity [13], and the exponential mechanism [12]. These techniques allow performing a wide scope of analyses in a differentially private manner, including conducting surveys over sensitive information, computing statistics, datamining, and sanitization. The reader is referred to [3] for a recent survey.

An immediate consequence of differential privacy is that (unless computing a constant function) a mechanism cannot compute a deterministic function. In other words, a differentially private version of an analysis would be a randomized approximation to the analysis, and furthermore, it would generally be possible to choose from a host of implementations for a task (e.g., the three generic construction techniques mentioned about may result with different mechanisms). Naturally, the designer of the analysis should choose one that is *useful*. Usefulness, however, depends on how the outcome of the analysis would be used, i.e., on the preferences of its consumer, that we henceforth refer to as *information consumer*. Such a trade-off between uncertainty and utility, while taking consumer's preferences into account, is the subject of rational-choice theory and decision theory, as noted in [9], [10].

We discuss the two models of utility which were previously discussed in [9], [10]. In both, the information consumer has *side information* (her own world-view or previous knowledge), and a *loss-function* which quantifies the consumer's preferences and the quality of the solution for her problem. Intuitively, it describes how bad is a deviation from the exact answer for the consumer, a measure of her intolerance towards the inaccuracy imposed by differentially private mechanisms. Finally, the models assume that the consumers are *rational* - they combine the structure of the mechanism, their side information and their personal loss-function (preferences) with the goal of minimizing their loss, or, equivalently maximizing their utility. The two models differ in the way side information is formulated and respectively how utility function is defined.

Information consumers' accuracy requirements vary: for some consumers only an exact answer would be of value, whereas others may aim at minimizing the estimate bias (ℓ_1 error), or its variance (ℓ_2 error), and, clearly, many other criteria exist. It seems that a discussion of the utility of differentially private mechanisms should take this rich variety into account. The recent work of Ghosh, Roughgarden, and Sundararajan [9] has put forward a serious attempt at doing exactly that with respect to (oblivious) Bayesian information consumers. In this utility model, the consumer's side information is described as an a priori

distribution on the exact result of the analysis. The recent work of Gupte and Sundararajan [10] considers a related model where the information consumers are *risk-averse*. Here, the information consumer's knowledge is a set of possible values the exact analysis can take, and an optimal mechanism minimizes the consumer's worst-case expected loss.

Composition theorems for differential privacy only guarantee that the degradation in privacy is not more than exponential in the number invocations. Hence, while different consumers may exhibit different optimal mechanisms, a very important goal is to avoid invoking that multiplicity of mechanisms. This degradation is part of the motivation for the work on *sanitization* where a family of queries are answered at once [5], [1], [8], [7], the work on *privacy under continual observation* [4], and the construction of the *Median Mechanism* [14]. A surprising result of Ghosh, Roughdarden, and Sundararajan [9] is that invoking a multiplicity of optimal mechanisms may not be necessary. They consider a database that is a collection of Binary inputs (e.g., pertaining to having some disease) and Bayesian information consumers that wish to count the number of *one* entries in the database (equivalently, compute the sum of the entries). They show the existence of a single mechanism that enables optimality for *all* Bayesian information consumers (the mechanism needs to be invoked only once). The mechanism itself is not optimal for all Bayesian information consumers, however, each consumer can perform a deterministic remapping on the outcome of the common mechanism, where the remapping is chosen according to her notion of utility, and locally output a result that is effectively according to one of her optimal mechanism. Such a common mechanism is referred to as *universally optimal*. An analogous result for risk-averse information consumers was shown in [10].

Are these results of [9] and [10] that deal with the simple case of a single count query "accidental", or can they be extended to other queries? to multiple queries? One would anticipate that universally-optimal mechanisms should exist (at least) for those queries that are closely related to counting, such as sum queries where the inputs are non-binary, histograms, and bundles of two or more count queries.

A. Our Results and Directions for Future Progress

In contrast with the anticipation expressed in the previous paragraph, we show that settings in which universally optimal mechanisms exist are extremely rare, and, in particular, in both the setting of Bayesian and of risk-averse information consumers, universally optimal mechanisms do not exist even for sum queries where the inputs are non-binary, histograms, and bundles of two or more count queries.

Moreover, in the case of Bayesian information consumers, we give a characterization of those functions of the data that admit universally optimal mechanisms. The characterization makes use of a combinatorial structure of the query function $f : \mathcal{D}^n \to \mathcal{R}_f$, where $\mathcal{D}$ is the domain of the database records and $\mathcal{R}_f$ is the output space of the query function. We define this combinatorial structure of the query G_f and call it a *privacy constraint graph*. The vertices of G_f correspond to values in $\mathcal{R}_f$, and edges correspond to pairs of values resulting by applying f to neighboring databases. (This graph was examined in some proofs in [11] as well). We show:

Theorem 14 (Informal)**.** *If G_f contains a cycle then no universally optimal mechanism exists for f.*

Theorem 15 (Informal)**.** *If G_f is a tree that contains a vertex of degree 3 or more, then no universally optimal mechanism exists for f for better values of the privacy parameter.*

Facing the impossibility of universal optimality, an alternative may be found in an approximate notion, which enables (approximate) optimality to (approximately) all of the information consumers. A good notion of approximate optimality should allow constructing such mechanisms for sum queries, histograms, and more. Furthermore, it should allow performing several queries and satisfy a composition requirement, in a sense that when applying two such mechanisms to two different queries, the resulting composed mechanism should be somewhat approximately optimal for the two queries together.

Finally, we note that, following prior work we focus on *oblivious* mechanisms (see Section II-B for the technical definition). In Section III, we show that for the intuitive generalizations of count queries, enabling *non*-oblivious universal mechanisms from which optimal oblivious mechanisms are derived, still leaves the construction of universally optimal mechanisms impossible. The question whether non-oblivious universally-optimal mechanisms exist for some other natural abstract queries, from which all oblivious universally-optimal mechanisms may be derived is left open.

B. Related Work

Most relevant to our work are the papers by Ghosh, Roughgarden, and Sundararajan [9] and by Gupte and Sundararajan [10]. Ghosh, Roughgarden and Sundararajan show that the geometric mechanism (a discrete version of the Laplace mechanism of [6]) yields optimal utility for all Bayesian information consumers for a count query. Their proof begins by observing that all differentially private mechanisms correspond to the feasible region of a Linear Program (a polytope), and that minimizing disutility can be expressed as minimizing a linear functional. Hence, every Bayesian information consumer has an optimal mechanism corresponding to a vertex of the polytope, which in turn corresponds to a subset of the constraints of the Linear Program which are tight (optimal mechanisms, not corresponding to the polytope vertices, may also exist). They introduce a *constraint matrix* that uniquely corresponds to a

vertex of the polytope, and indicates which constraints are tight, and which are slack on that vertex. Those constraint matrices that correspond to optimal mechanisms, are shown to have some special structure that allows to derive mechanisms with the same signature (and thus equal) from the geometric mechanism using some deterministic remapping on its output.

We are also interested in observing the tight constraints in some mechanisms. We will not need the full description of the structure of such a constraint matrix. Instead we only use the observation that tight privacy constraints can be derived only from mechanisms that also obey similar tight constraints.

Gupte and Sundararajan show similar results for the risk-averse utility model, where consumers try to minimizes their maximal worst-case disutility. They provide a full characterization of the mechanisms which are derivable (by random remapping) from the geometric mechanism and use this characterization to construct a universally-optimal mechanism for a count query. An interesting feature of the construction is that it releases noisy answers of the query at different privacy levels, thus keeping more privacy against specific consumers, and enabling more utility to others.

Also related to our work is the recent work of Kifer and Lin [11] that studies privacy and utility, in a very general setting, from an axiomatic point of view. They introduce a partial order on mechanism where mechanism Y is at least as *general* as mechanism X if X can be derived from Y by post processing. They also introduce the concept of maximal generality, which turns to be useful in our proofs.

II. Preliminaries

A. Differential Privacy [6]

Consider databases $D_1, D_2 \in \mathcal{D}^n$ which consist of n records out of some domain $\mathcal{D}$. The Hamming Distance between D_1 and D_2 is the number of records on which they differ. We will call databases at distance one *neighboring*.

Definition 1 (Differential Privacy [6])**.** Let $\mathcal{M} : \mathcal{D}^n \to \mathcal{R}$ be a probabilistic mechanism. $\mathcal{M}$ preserves *α-differential-privacy* for $\alpha \in (0,1)$ if for any two neighboring databases $D_1, D_2 \in \mathcal{D}^n$ and any (measurable) subset of the mechanism's range $S \subseteq \mathcal{R}$,

$$Pr\left[\mathcal{M}(D_1) \in S\right] \geq \alpha \cdot Pr\left[\mathcal{M}(D_2) \in S\right]. \qquad (1)$$

The probability is taken over the coin tosses of the mechanism $\mathcal{M}$.

Notice that the greater α is the less the mechanism's output depends on the exact query result, and so better privacy is attained.

B. Oblivious Mechanisms

We consider a setting where several information consumers are interested in estimating the value of some query $f(\cdot)$ applied to a database $D \in \mathcal{D}^n$, and answered by a differentially private mechanism $\mathcal{M}$. Reference [9] shows that if no restriction is put on the mechanism, then no universally optimal mechanism exists for count queries (intuitively, universal optimality, defined below, means that all potential consumers minimize their loss simultaneously). On the other hand, universally optimal mechanisms sometimes do exist if we restrict our mechanisms such that their output distribution depends only on the the exact query result (a.k.a. *oblivious mechanisms*). This is why in [9] (and later in [10]) only oblivious mechanisms are considered[1]. We follow suit and only consider oblivious mechanisms. We show in Subsection III-B that this restriction does not weaken the basic results presented in Section III.

Definition 2 (Oblivious Mechanism)**.** Let $f : \mathcal{D}^n \to \mathcal{R}_f$ be a query. A mechanism $\mathcal{M} : \mathcal{D}^n \to \mathcal{R}$ is *f-oblivious* (or simply *oblivious*) if there exists a randomized function $\tilde{\mathcal{M}} : \mathcal{R}_f \to \mathcal{R}$ such that, for all $D \in \mathcal{D}^n$, the distributions induced by $\mathcal{M}(D)$ and $\tilde{\mathcal{M}}(f(D))$ are identical.

Combining α-differential privacy with obliviousness, we get that for every $i, i' \in \mathcal{R}_f$ which are outputs of neighboring databases D, D' (i.e., $f(D) = i$ and $f(D') = i'$), then $\Pr[\tilde{\mathcal{M}}(i) \in S] \geq \alpha \cdot \Pr[\tilde{\mathcal{M}}(i') \in S]$ for all $S \subseteq \mathcal{R}$.

Oblivious Differentially Private Mechanisms for a Count Query: An oblivious finite-range mechanism $\mathcal{M} : \mathcal{D}^n \to \mathcal{R}$ estimating $f : \mathcal{D}^n \to \mathcal{R}_f$ can be described by a row-stochastic matrix $X = (x_{i,j})$ of the underlying randomized mapping $\tilde{\mathcal{M}}$, whose rows are indexed by elements of $\mathcal{R}_f$, and whose columns are indexed by elements of $\mathcal{R}$, where $x_{i,j}$ equals the probability of outputting $j \in \mathcal{R}$ when $f(D) = i$. Since $\mathcal{R}$ is finite, and information consumers anyway remap the outcome of $\mathcal{M}$, we can assume, wlog, that $\mathcal{R} = \{0, 1, 2, \ldots, |\mathcal{R}| - 1\}$.

We now consider the case where $\mathcal{D} = \{0, 1\}$ and $f(D)$ counts the number of one entries in D. Hence, $\mathcal{R}_f = \{0, \ldots, n\}$ and the matrix X is of dimensions $(n+1) \times |\mathcal{R}|$. Preserving α-differential privacy poses constraints on the transition matrix X beyond row-stochasticity. Note that for the count query, the query results of two neighboring databases may differ by at most one. Differential privacy hence imposes the constrains $x_{i,j} \geq \alpha \cdot x_{i+1,j}$ and $x_{i+1,j} \geq \alpha \cdot x_{i,j}$ where $i \in \mathcal{R}_f = \{0 \ldots n-1\}$ and $j \in \mathcal{R}$. Adding row-stochasticity and differential privacy, we get that an oblivious differentially private mechanism for the count query should satisfy the following linear constraints:

$$x_{i,r} \geq \alpha x_{i+1,r} \qquad \forall i \in \{0, \ldots, n-1\}, \forall r \in \mathcal{R} \qquad (2)$$

$$\alpha x_{i,r} \leq x_{i+1,r} \qquad \forall i \in \{0, \ldots, n-1\}, \forall r \in \mathcal{R} \qquad (3)$$

[1]Impossibility of universal optimality when the mechanisms are not restricted to being oblivious is proved in [9] for Bayesian information consumers. For risk-averse consumers, [10] show that non-oblivious mechanisms may be replaced with oblivious ones without affecting the consumers' utility for the worse.

$$\sum_{r\in\mathcal{R}} x_{i,r} = 1 \qquad \forall i \in \{0,\ldots,n\} \tag{4}$$

$$x_{i,r} \geq 0 \qquad \forall i \in \{0,\ldots,n\}, \forall r \in \mathcal{R} \tag{5}$$

C. Utility Models

We use the utility models defined in [9] and [10]. In both, a *loss function* $\ell(i,r)$ quantifies an information consumer's disutility when she chooses to use answer r while the correct answer is i. Given a loss function $\ell(\cdot,\cdot)$ of an information consumer, if the exact answer is i then her expected loss is $\sum_{r\in\mathcal{R}} x_{i,r}\cdot\ell(i,r)$.[2] Loss functions vary between consumers, and the only assumptions made in [9], [10] is that $\ell(i,r)$ depends on i and $|i-r|$ and is monotonically non-decreasing in $|i-r|$ for all i. (This is a reasonable requirement that turns to be crucial for the existence of a universally optimal mechanism [9].) Examples of loss functions include $\ell_1(i,r) = |i-r|$ (consumers who care to minimize expected mean error); $\ell_2(i,r) = (i-r)^2$ (minimize error variance); and $\ell_{bin}(i,r)$ that evaluates to 0 if $i = r$ and to 1 otherwise (minimize number of errors).

Information consumers differ in their knowledge about the exact $f(D)$. References [9] and [10] model this knowledge differently as we now describe.

Bayesian Model [9]: In the Bayesian utility model, an information consumer's knowledge is represented by a vector $\bar{p}$ where p_i is the consumer's a priori probability that $f(D) = i$. Having a vector of prior probabilities $\bar{p}$ and loss function $\ell(\cdot,\cdot)$, the consumer's expected loss can be expressed as $\sum_i p_i\cdot\sum_r x_{i,r}\cdot\ell(i,r)$. The *optimal mechanisms* for this information consumer hence are the solutions of the linear program in the variables $x_{i,r}$ consisting the constraints in (2)–(5) and the objective

$$\text{minimize} \sum_{i\in\mathcal{R}_f} p_i \cdot \sum_{r\in\mathcal{R}} x_{i,r}\cdot\ell(i,r). \tag{6}$$

Risk-Averse Model [10]: In the risk-averse utility model an information consumer's knowledge restricts the possible values for the exact $f(D)$. This is expressed by a set $S \subseteq \mathcal{R}_f$ of the possible values $f(D)$ can take. The consumer is interested in minimizing her maximal expected loss conditioned on $f(D) \in S$, i.e., $\max_{i\in S}\sum_r x_{i,r}\cdot\ell(i,r)$. Similarly to the above, the optimal mechanism for an information consumer is a solution to a linear program consisting the constraints in (2)–(5) and the objective

$$\text{minimize} \max_{i\in S} \sum_{r\in\mathcal{R}} x_{i,r}\cdot\ell(i,r). \tag{7}$$

D. Remapping and Generality

An information consumer might have access to a private mechanism U which is not tailored specifically for her needs (i.e., to her prior knowledge and loss function). Yet, she may be able to recover a better mechanism for her needs by means of post-processing, which we will denote *remapping*. To intuit remapping, consider a consumer that knows that for the specific database the count query cannot yield the answer 0. If that consumer receives a 0, it may be beneficial for her to remap it to 1. (Recall that the loss function is monotone in $|i-r|$.) Denoting the given mechanism by U and the remapping by T (a row-stochastic linear transformation, T has no access to the information of the database other then the output of U), the actual mechanism that is used by the information consumer is $T\circ U$ (in matrix form: UT).

Notice that given a mechanism U with a finite range, an information consumer can find the optimal remapping T for her (such that $T\circ U$ has optimal utility), by constructing a linear program in which $T = (t_{i,j})$ are the program variables [10].

Definition 3 (Derivable Mechanisms, Generality Partial Order [11])**.** Let X, Y be private mechanisms. We say that a mechanism X is *derivable* from a mechanism Y if there exists a random remapping T of the results of mechanism Y, such that $X = T\circ Y$. We also say that Y is *at least as general as* X, and denote this relation by $X \preceq_G Y$. If $X \preceq_G Y$ and $Y \preceq_G X$ we say that X, Y are *equivalent*.

Definition 4 (Maximal Generality [11])**.** Let X be an α-differentially private mechanism. X is *maximally general* if for every α-differentially private mechanism Y, if $X \preceq_G Y$ then $Y \preceq_G X$.

After introducing the notion of maximally general mechanisms (for any definition of privacy), Kifer and Lin fully characterize all maximally general private mechanisms with a finite input space in the differential privacy setting. First they introduce the concept of *column-graphs*[3] of a private mechanism, which mark the tight privacy constraints in one column of the mechanism X.

Definition 5 (Column graph [11])**.** Let X be an α-differentially private mechanism with a finite input space. Let r be some possible output of X, and x_r be its corresponding column in X. Let I be the input space of X (corresponding to X's rows). The graph associated with this column has I as the set of nodes, and for any $i_1, i_2 \in I$, there is a directed edge (i_1, i_2) if i_1 and i_2 match neighboring databases and $x_{i_1,r} = \alpha x_{i_2,r}$, and a directed edge (i_2, i_1) if $x_{i_2,r} = \alpha x_{i_1,r}$. The direction of the edges is only necessary to distinguish between maximally general mechanisms which have similar undirected column-graphs, but it will not be essential to the rest of this article.

Kifer and Lin characterize the maximally general differ-

[2] This is only true if the consumer uses the mechanism X *directly*, i.e., the consumer leaves the mechanism's output as is, and does not apply a post-processing step. The ability to apply such a post-processing step on the mechanism's output will be discussed in the next sub-section.

[3] Kifer and Lin actually define *row graphs* and not *column graphs*. We follow the matrix structure of [9], [10] which is simply the transposed matrix of the one used by Kifer and Lin, hence the difference in terminology.

entially private mechanisms with a finite input space:

Theorem 6 ([11])**.** *Fix a privacy parameter α and a database query f with a finite range for databases of a specific size. Let X be an α-differentially private mechanism with a finite range. Then X is maximally general iff each column graph of X's columns (according to the privacy constraints implied by f) is connected.*

This theorem shows that we wish to maximize the set of tight privacy constraints in order to make a private mechanism as general as possible. Notice that having just one entry of a column in X and the spanning tree of this column's graph (we need to know the direction of the edges as well), determines all the entries of this column.

E. Universal Mechanisms

Consider a collection of Bayesian information consumers, and suppose we wish to enable each of the information consumers to sample a result from a differentially private mechanism optimizing her utility. In [9], Ghosh, Roughgarden and Sundararajan showed that this does not necessarily require executing multiple mechanisms: if the query is a count query, then it is possible to construct one *universally optimal* mechanism U, from which all information consumers can *simultaneously* recover an optimal mechanism for their needs by *remapping*. I.e., every information consumer has an optimal private mechanism which is derivable from U. This result is repeated for risk-averse information consumers by Gupte and Sundararajan in [10]. More formally:

Theorem 7 (Universal optimality, Bayesian consumers [9])**.** *Fix a privacy parameter $\alpha \in (0,1)$. There exists an α-differentially private mechanism U for a single count query, such that for every prior $\bar{p}$ and every monotone loss function $\ell(\cdot,\cdot)$ there exists a (deterministic) remapping T such that $T \circ U$ implements an optimal oblivious mechanism for $\bar{p}, \ell(\cdot,\cdot)$.*

Theorem 8 (Universal optimality, risk-averse consumers [10])**.** *Fix a privacy parameter $\alpha \in (0,1)$. There exists an α-differentially private mechanism U for a single count query, such that for every set S of possible outcomes and every monotone loss function $\ell(\cdot,\cdot)$ there exists a (probabilistic) remapping T such that $T \circ U$ implements an optimal oblivious mechanism for $S, \ell(\cdot,\cdot)$.*

In both theorems U is realized by the geometric mechanism (a variant of adding Laplace noise of [6]). It is shown that for every information consumer there is at least one private mechanism that is derivable from the geometric mechanism and is optimal for her.

III. Impossibility of Universally Optimal Mechanisms for Generalizations of Count

When the domain of the database records is $\{0,1\}$, a count query is equivalent to a sum query. Theorems 7 and 8 can hence be thought of as applying to a *sum query* over the integers, where the domain of the database is Binary. It is natural to ask whether the results of these theorems can be extended to showing that universally optimal mechanisms exist for sum queries when the underlying data is taken from a larger domain such as $\mathcal{D} = \{0,1,\ldots,m\}$ where $m \geq 2$. We answer this question negatively.

Consider the case $m = 2$. Recall that an oblivious differentially private mechanism can be described by a row-stochastic matrix $X = (x_{i,j})$, such that $x_{i,j}$ is the probability of the mechanism to return j when the exact result is i. A difference of the case $m = 2$ from count queries ($m = 1$) is that applying a sum query to two neighboring databases may yield results which differ by 0, 1, or 2 (instead of 0 or 1). Therefore, in the linear program describing mechanism X, (2) and (3) should be replaced by the following four constraints: $x_{i,r} \geq \alpha x_{i+1,r}$, $\alpha x_{i,r} \leq x_{i+1,r}$ for all $0 \leq i \leq 2n-1$ and $r \in \mathcal{R}$; and $x_{i,r} \geq \alpha x_{i+2,r}$, $\alpha x_{i,r} \leq x_{i+2,r}$ for all $0 \leq i \leq 2n-2$ and $r \in \mathcal{R}$. (The range for i in the other equations should be modified to $0,\ldots,2n$.)

Once again, a consumer's optimal mechanism can be found by solving a linear program with all the constraints and the appropriate target function.

A. The Basic Impossibility Result for Sum Queries

We first consider the case where the database contains $n = 1$ record, taking values in {0,1,2} (i.e., $m = 2$). Later, we generalize to $n \geq 1$ and $m \geq 2$. Note that in the case of $n = 1$, the non-oblivious mechanisms are identical to oblivious mechanisms. We consider non-oblivious universal mechanisms as well when generalizing this result lo larger values of n.

Observation 9. *In the Bayesian model there exists an information consumer whose only optimal mechanism is $X = \frac{1}{1+2\alpha} \cdot \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix}$ and an information consumer whose optimal mechanisms are all of the form $Y = \frac{1}{1+\alpha} \cdot \begin{bmatrix} 1 & \alpha & 0 \\ \alpha & 1 & 0 \\ q & 1+\alpha-q & 0 \end{bmatrix}$, where $q \in [\alpha, 1]$.*

Proof: Consider an information consumer with a prior $\bar{p} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and a loss function ℓ_{bin} (i.e., a penalty of 1 whenever she chooses an answer different from the exact result, and no penalty otherwise). It is easy to see that no optimal mechanism for this consumer outputs a value not in $\{0,1,2\}$.

The information consumer wishes to minimize

$$\sum_{i=0}^{2} p_i \sum_{r=0}^{2} x_{i,r} \cdot \ell(i,r) = \frac{1}{3} \sum_{i=0}^{2} \sum_{r \neq i} x_{i,r}$$
$$= \frac{1}{3} \sum_{i=0}^{2} (1 - x_{i,i}) = 1 - \frac{1}{3} \sum_{i=0}^{2} x_{i,i}.$$

And so, the consumer's goal is to maximize $\sum_{i=0}^{2} x_{i,i}$ subject to maintaining α-differential privacy.

For $i \in \{0,1,2\}$, having α-differential privacy implies

$$\alpha x_{i,i} \leq x_{j,i} \quad \forall j \in \{0,1,2\} \setminus \{i\}, \tag{8}$$

and hence (by summing up (8) for $j \neq i$), we get

$$2\alpha x_{i,i} = \sum_{\substack{j=0 \\ j\neq i}}^{2} \alpha x_{i,i} \leq \sum_{\substack{j=0 \\ j\neq i}}^{2} x_{j,i}. \tag{9}$$

Summing up (9) for $i \in \{0,1,2\}$ we get

$$\sum_{i=0}^{2} 2\alpha x_{i,i} \leq \sum_{i=0}^{2} \sum_{\substack{j=0 \\ j\neq i}}^{2} x_{j,i} = \sum_{i=0}^{2} (1 - x_{i,i}) = 3 - \sum_{i=0}^{2} x_{i,i},$$

and we can now conclude that $\sum_{i=0}^{2} x_{i,i} \leq \frac{3}{2\alpha+1}$. This inequality is tight iff (8) is tight (i.e., $x_{j,i} = \alpha x_{i,i}$) for every $i \neq j$. In that case, we get the following system of linear equations:

$$\begin{aligned} x_{11} + \alpha x_{22} + \alpha x_{33} &= 1 \\ \alpha x_{11} + x_{22} + \alpha x_{33} &= 1 \\ \alpha x_{11} + \alpha x_{22} + x_{33} &= 1 \end{aligned}$$

Since the three equations are linearly independent, we get a *unique* solution: $x_{1,1} = x_{2,2} = x_{3,3} = \frac{1}{1+2\alpha}$.

A similar proof shows that mechanisms of the form Y are the only mechanisms optimal for information consumers with a prior $p_0 = p_1 = \frac{1}{2}, p_2 = 0$ and loss function ℓ_{bin}.

It may seem like we restrict ourselves only to information consumers with the ℓ_{bin} loss function. Note that, according to Theorem 6, there are not so many maximally general mechanisms whose range is a subset of $\{0,1,2\}$, and some of them are not optimal for any consumer. Therefore, the mechanisms described are also the only optimal mechanisms for a variety of other information consumers, such as whose prior is $p_0 = p_1 = \frac{1}{2}, p_2 = 0$ and loss function is ℓ_1. Also, even more such consumers can be found easily in any sequence of consumers which converge to consumers with such unique optimal mechanisms (i.e., their priors and loss functions converge to the prior and loss function of the consumer we chose). Such information consumers with close priors and close loss functions to the ones described above will have the same unique optimal mechanisms. ■

Observation 10. *In the risk-averse model there exists an information consumer whose only optimal mechanism is* $X = \frac{1}{1+2\alpha} \cdot \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix}$ *and an information consumer whose optimal mechanisms are all of the form* $Y = \frac{1}{1+\alpha} \cdot \begin{bmatrix} 1 & \alpha & 0 \\ \alpha & 1 & 0 \\ q & 1+\alpha-q & 0 \end{bmatrix}$, *where* $q \in [\alpha, 1]$.

Proof: Consider an information consumer whose loss function is ℓ_{bin} who knows the support of the query is $S = \{0,1,2\}$. As in the previous observation, the support of any optimal mechanism for this consumer must be a subset of $\{0,1,2\}$. Notice that if the consumer uses the mechanism described by X then her maximal expected loss is $\frac{2\alpha}{1+2\alpha}$.

Assume for a contradiction that the consumer has another mechanism X' with maximal expected loss at most $\frac{2\alpha}{1+2\alpha}$. I.e.,

$$\max\{x'_{0,1}+x'_{0,2}, x'_{1,0}+x'_{1,2}, x'_{2,0}+x'_{2,1}\} \leq \frac{2\alpha}{(1+2\alpha)}. \tag{10}$$

Since $X' \neq X$, (10) implies that $x'_{i,j} < \frac{\alpha}{1+2\alpha}$ for some $i \neq j$. Taking into account that X' is α-differentially private we get $x'_{j,j} \leq \frac{1}{\alpha} \cdot x'_{i,j} < \frac{1}{1+2\alpha}$, and hence the maximal expected loss is at least $\sum_{i\neq j} x'_{i,j} = 1 - x'_{j,j} > 1 - \frac{1}{1+2\alpha} = \frac{2\alpha}{1+2\alpha}$, in contradiction to the assumption that this mechanism is at least as good as X for this information consumer.

A similar proof shows that mechanisms of the form Y are the only mechanisms optimal for an information consumer with auxiliary knowledge of the support $S = \{0,1\}$ and loss function ℓ_{bin}. As in the previous observation, the mechanisms described are also the only optimal mechanisms for a variety of other information consumers. ■

We will now use these two observations to show that in both models no universally optimal mechanism U exists. (This is true even if we allow U to have a non-discrete range.)

Claim 11. *No α-differentially private mechanism can derive both X and an instance of Y.*

Proof: Assume for a contradiction that such a mechanism U exists, so X and some instance of Y are both derivable from U. For simplicity we refer to this instance as Y. By Theorem 6, X is a maximally general mechanism. Therefore $U \preceq_G X$, and hence $Y \preceq_G X$, i.e., there exists a random remapping T such that $Y = XT$. Denote by x_j the j^{th} column of X, and by y_k the k^{th} column of Y. We get that $y_k = t_{0,k} \cdot x_0 + t_{1,k} \cdot x_1 + t_{2,k} \cdot x_2$ for $k \in \{0,1,2\}$. Note that some α-differentially privacy constraints in Y are tight. Specifically, $y_{1,0} = \alpha y_{0,0}$ and $y_{0,1} = \alpha y_{1,1}$. As Y's columns are non-negative linear combinations of X's columns, such a tight constraint in a column of Y appears only if this column is a linear combination of columns of X in which the same privacy constraints are also tight. Note that the first two entries of every column in Y correspond to a tight constraint. But since $x_{0,2} = x_{1,2} > 0$, mapping this column of X by T to any column of Y (even with just a positive probability), yields a mechanism with a column in which the first two entries do not correspond to a tight constraint. Therefore, a contradiction. ■

B. Generalizing the Basic Result for Other Queries

Next, we generalize the basic impossibility result to the more general case of sum queries where $m \geq 2$ and $n \geq 1$. Another natural generalization of count queries is to histogram queries which partition the database records into

three categories or more. Finally we consider the generalization of single count queries to a bundle of count queries, where a bundle contains several simple (non-trivial) count queries that need to be answered simultaneously. Note that a consumer's disutility for a bundle query need not be the sum of the losses for the separate basic queries – it may be a more involved function of the bundle outputs. Furthermore, information consumers may have auxiliary knowledge about the dependency between bundle outputs.

Theorem 12. *No universally optimal mechanism exists for sum queries for databases whose records take values in the set $\{0, 1, \ldots, m\}$ where $m \geq 2$. No universally optimal mechanism exists for histogram queries, except for histograms for one predicate and its complement or trivial predicates. No universally optimal mechanism exists for bundles of more than one simultaneous non-trivial count queries. These results hold both for the Bayesian and the risk-averse utility models.*

All the generalization queries which were mentioned in the theorem have a common feature which distinguishes them from count queries. In all the suggested queries there are 3 (or more) query outputs which are the results of 3 neighboring databases. This fact enables us to reduce this problem to the basic case of the previous subsection. The proof is omitted and will appear in the full version of this article. We note that this statement holds even if we allow *non*-oblivious universal mechanisms.

IV. A Characterization of Universal Optimality in the Bayesian Setting

We now discuss a more general setting, where a query (not necessarily related to sum or count) is answered by a differentially private mechanism in the Bayesian utility model. We follow other works on this subject and only consider oblivious private mechanisms. Note that although our results do not exclude the possibility of non-oblivious differentially private mechanisms, our techniques yield that no such non-oblivious universally optimal mechanisms exist for many natural functions. Specifically, enabling universal non-oblivious mechanisms cannot resolve such impossibilities for a query whenever there are 3 (or more) values which are the exact query results of 3 different neighboring databases. This is due to the same argument that was used in Subsection III-B.

Let the database records be taken from a discrete domain $\mathcal{D}$ and let the query be $f : \mathcal{D}^n \to \mathcal{R}_f$ (wlog, we will assume that f is a surjective function, in which case $\mathcal{R}_f = \{f(D) : D \in \mathcal{D}^n\}$ is also a discrete set). Define the following graph where edges correspond to answers f may give on neighboring databases (and hence to restrictions on output distributions implied by differential privacy):

Definition 13 (Privacy Constraint Graph)**.** Fix a query $f : \mathcal{D}^n \to \mathcal{R}_f$. The *Privacy Constraint Graph* for f is the undirected graph $G_f = (V, E)$ where $V = \mathcal{R}_f$ is the set of all possible query results and $E = \{(f(D_1), f(D_2)) : D_1, D_2 \in \mathcal{D}^n \text{ are neighboring}\}$. The *degree* of the constraint graph, $\Delta(G_f)$, is the maximum over its vertices' degrees. For $i_2, i_2 \in \mathcal{R}_f$, G_f induces a distance metric $d_{G_f}(i_1, i_2)$ that equals the length of the shortest path in G_f from i_1 to i_2.

Observe that the constraint graph is connected for any query f: If $i_1 = f(D_1)$ and $i_2 = f(D_2)$ then there is a sequence of neighboring databases starting with D_1 and ending in D_2, and hence a path from i_1 to i_2 in G_f.

Recall that the results of [9], [10] are restricted to loss functions $\ell(i, r)$ that are monotonically non-decreasing in the metric $|i - r|$. In our more general setting, we avoid interpreting outcome of f as points of a specific metric space, and hence we only consider the ℓ_{bin} loss function, which would remain monotone under any imposed metric.

Outline of this Section.: We are now ready to describe the results of this section. Let f be a query, and G_f its constraint graph. We first show that if G_f is a single cycle, then no universally optimal mechanism exists for f. This impossibility result is then extended to the case where G_f contains a cycle.

Theorem 14. *Fix a query $f : \mathcal{D}^n \to \mathcal{R}_f$, and let G_f be its constraint graph. Consider Bayesian information consumers with loss function ℓ_{bin}. If G_f contains a cycle then no universally optimal mechanism exists for these consumers.*

Constraint graphs of sum queries (for $m \geq 2$), histograms and bundles of queries all have cycles of length 3, so, in the Bayesian utility model, Theorem 14 generalizes all our previous results.

Next, we consider the case where G_f is a tree and show that if G_f contains a vertex of degree 3 or higher, then no α-differentially private universally optimal mechanism exists for f for $\alpha > 1/(\Delta(G_f) - 1)$. (Recall that the closer α is to one, the better privacy we get.)

Theorem 15. *Fix a query $f : \mathcal{D}^n \to \mathcal{R}_f$, and let G_f be its constraint graph. Consider Bayesian information consumers with loss function ℓ_{bin}. If the privacy parameter $\alpha > 1/(\Delta(G_f) - 1)$ then no universally optimal mechanism exists for these consumers.*

We can conclude from theorems 14 and 15 that for $\alpha > 0.5$, the only functions f for which universally optimal mechanisms exist are those where G_f is a simple chain, as is the case for the count query.

The proof structure is similar to the one presented in the previous section for sum queries. We begin with the case where G_f is a simple cycle. We consider two consumers with different priors and loss function ℓ_{bin}, and show that the optimal mechanisms for these consumers must have specific structures (in the sense that some privacy constraints

are satisfied tightly). Once again, we show that for two mechanisms with such structures, there is no mechanism which is at least as general as these two (i.e., there is no single mechanism which derives both of them).

Next, we extend the proof to the case where G_f contains a cycle. We focus on a cycle in G_f of smallest size m, and consider two information consumers. The consumers are similar to those for the case where G_f is a cycle, and so are the optimal mechanisms for them, except that we need to prove that these optimal mechanisms can be extended in a differentially private manner to the entire range of f. For that we introduce a labeling of G_f in which the labels of adjacent vertices differ by at most one modulo m.

Last, we discuss the case where G_f is a tree containing a vertex of degree at least 3. Focusing on that vertex and three of its adjacent vertices, we present three consumers with different priors. Again, we focus on the corresponding entries in the matrices of their optimal mechanisms, and find which constraints must be tight. Assuming all three mechanisms are derived from a single mechanism U, we present three different partitions of U's range according to which constraints are tight for every measurable subset of U's range. Combining the attributes from these partitions, we get one elaborated partition of U's range. We can then assume U's range is finite and reveal the structure of its matrix columns. Such a structure of U's columns (for the consumers we chose) is feasible iff we compromise for a privacy parameter $\alpha \leq 0.5$. Finally, we generalize this claim to any degree of one vertex.

A. The Basic Case: G_f is a Cycle

We begin with the simple case where G_f is a single cycle of $m > 2$ vertices[4].

Claim 16. *If the constraint graph G_f of $f : \mathcal{D}^n \to \mathcal{R}_f$ is a single cycle, then no universally optimal mechanism for Bayesian information consumers exists for f.*

Proof: Assume G_f is the cycle $C_m = (v_0, v_1, \ldots, v_{m-1}, v_0)$. We already proved impossibility of universal optimality for the case $m = 3$ in Claim 11. We now deal with the case $m > 3$. As in the proof of Claim 11, we will present two information consumers, and their corresponding optimal mechanisms, and prove that these cannot be derived from a single mechanism.

We first consider an information consumer with loss function ℓ_{bin} and prior $p_{v_0} = p_{v_1} = \cdots = p_{v_{m-1}} = 1/m$, and construct the unique optimal mechanism X for this consumer. (X is represented by an $m \times m$ matrix since with the ℓ_{bin} loss function the support of the optimal mechanism's range must match the support of the consumer's prior.) An optimal mechanism minimizes

$$\sum_{v_i \in C_m} p_{v_i} \sum_{r \in C_m} x_{v_i,r} \cdot \ell_{bin}(v_i, r) = \sum_{v_i \in C_m} p_{v_i} \cdot (1 - x_{v_i,v_i}) = 1 - \frac{1}{m} \sum_{v_i \in C_m} x_{v_i,v_i},$$

and hence, the consumer's goal is to maximize $\sum_{v_i \in C_m} x_{v_i,v_i}$ subject to maintaining α-differential privacy. Maintaining α-differential privacy implies

$$\alpha^{d_{G_f}(v_i,v_j)} x_{v_i,v_i} \leq x_{v_j,v_i} \quad \forall v_i, v_j \in C_m, \tag{11}$$

and hence, by summing up the inequalities for all v_i, v_j, we get

$$\sum_{v_i \in C_m} \sum_{\substack{v_j \in C_m \\ v_j \neq v_i}} \alpha^{d_{G_f}(v_i,v_j)} x_{v_i,v_i} \leq \sum_{v_i \in C_m} \sum_{\substack{v_j \in C_m \\ v_j \neq v_i}} x_{v_j,v_i} = \sum_{v_i \in C_m} 1 - x_{v_i,v_i} = m - \sum_{v_i \in C_m} x_{v_i,v_i},$$

and we conclude that

$$\sum_{v_i \in C_m} x_{v_i,v_i} \leq \frac{m}{1 + \sum_{\substack{v_j \in C_m \\ v_i \neq v_j}} \alpha^{d_{G_f}(v_i,v_j)}}.$$

This inequality is tight iff (11) is tight (i.e., $x_{v_j,v_i} = \alpha^{d_{G_f}(v_i,v_j)} x_{v_i,v_i}$) for every $v_i \neq v_j \in C_m$. In such a case, we can find the mechanism's entries by solving a system of m linear equations (the sum of each row in the mechanism must be 1), in a similar argument to the one presented in the proof of Observation 9. Since these are m independent linear equations in m variables, our optimal solution for $x_{v_1,v_1}, \ldots, x_{v_m,v_m}$ is unique.

Utilizing the symmetry of the equations, we get that every row of X is a cyclic shift of:

$$\begin{aligned} &\delta \cdot (1, \alpha^1, \alpha^2, \ldots, \alpha^{(m-1)/2}, \alpha^{(m-1)/2}, \alpha^{(m-1)/2-1}, \ldots, \alpha^2, \alpha^1) \\ &\quad \text{if } m \text{ is odd,} \\ &\delta \cdot (1, \alpha^1, \alpha^2, \ldots, \alpha^{m/2-1}, \alpha^{m/2}, \alpha^{m/2-1}, \ldots, \alpha^2, \alpha^1) \\ &\quad \text{if } m \text{ is even.} \end{aligned} \tag{12}$$

where δ is chosen such that X is row-stochastic. The mechanism X satisfies α-differential privacy, it is optimal for our information consumer, and it is unique.

Our second information consumer uses ℓ_{bin} as her loss function, and prior $p_{v_0} = p_{v_1} = p_{v_2} = 1/3$ and $p_{v_3} = \cdots = p_{v_{m-1}} = 0$. Note that since $m > 3$ the vertices v_0, v_2 are not adjacent in G_f (so $d_{G_f}(v_0, v_2) = 2$). In constructing an optimal mechanism Y for the information consumer we will only consider the rows and columns pertaining to vertices v_0, v_1, v_2, noting that the columns for all other vertices contain only zeros, and there is some freedom with respect to the rows for the other vertices. Applying similar arguments as for mechanism X, we get that the columns of Y are of the forms $(1, \alpha^1, \alpha^2)^T$, $(\alpha^1, 1, \alpha^1)^T$, $(\alpha^2, \alpha^1, 1)^T$ (each of

[4] An example query that yields such a graph is $f : \{0,1\}^n \to [m]$ defined as $f(d_1, \ldots, d_n) = \sum_{i=1} d_i \mod m$. If $n \geq m > 2$ then G_f is a cycle of size m.

the columns may be multiplied by a different coefficient). By forcing row stochasticity, we can solve the following equations to get the coefficients:

$$\begin{bmatrix} 1 & \alpha^1 & \alpha^2 \\ \alpha^1 & 1 & \alpha^1 \\ \alpha^2 & \alpha^1 & 1 \end{bmatrix} \times \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

and we get a unique structure on the entries of these rows and columns of Y. This mechanism is of no surprise, as these entries are merely the finite-range version of the geometric mechanism (as shown in [9]).

Summarizing our findings, we get that

$$X = \delta \cdot \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^2 & \alpha \\ \alpha & 1 & \alpha & \cdots & \alpha^3 & \alpha^2 \\ \alpha^2 & \alpha & 1 & \cdots & \alpha^4 & \alpha^3 \\ \vdots & \vdots & & \ddots & & \vdots \\ \alpha & \alpha^2 & \alpha^3 & \cdots & \alpha & 1 \end{bmatrix}; Y = \begin{bmatrix} c_1 & c_2 \cdot \alpha^1 & c_3 \cdot \alpha^2 & 0 & \cdots & 0 \\ c_1 \cdot \alpha^1 & c_2 & c_3 \cdot \alpha^1 & 0 & \cdots & 0 \\ c_1 \cdot \alpha^2 & c_2 \cdot \alpha^1 & c_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 & \cdots & 0 \end{bmatrix} \tag{13}$$

We now show that instances of such mechanisms X and Y are not derivable from a single mechanism. Since the conditions stated for these mechanisms are necessary for them to be optimal for the two consumers we chose, this will prove that there is no universally optimal mechanism in such a scenario.

Suppose, towards a contradiction, that there exists a mechanism U which derives both X and some instance of Y. According to the characterization of generally maximal differentially private mechanisms (Theorem 6), X is maximally general. Therefore, we get that U is derivable from X and so Y is derivable from X as well. Therefore, there exists a remapping matrix T such that $Y = XT$. Remember that Y's columns are linear combinations of X's columns with non-negative coefficients, as described in the proof of Claim 11. Any tight constraint met in one of Y's columns must match the same tight constraints in all of X's columns which appear in the linear combination of that column. Once again, any specific column of X must appear in at least one linear combination of one of Y's columns with a positive coefficient (as any possible output of X must be remapped to the values $\{v_0, v_1, v_2\}$ by T). Notice that one of X's columns is $\delta(\alpha^{(m-1)/2}, \alpha^{(m-1)/2}, \alpha^{(m-1)/2-1}, \ldots, 1, \ldots, \alpha^{(m-3)/2})^T$ if m is odd and otherwise it is $\delta(\alpha^{m/2-1}, \alpha^{m/2}, \alpha^{m/2-1}, \ldots, 1, \ldots, \alpha^{m/2-2})^T$. Mapping this column into any of Y's first three columns (with any positive probability) cannot yield the tight constraints which appear in the first three entries of the chosen column in Y. Therefore, no such remapping T is feasible and we get a contradiction. ■

B. Impossibility of Universal Optimality When G_f Contains a Cycle

We now give a proof for Theorem 14 which deals with the case where G_f *contains* a cycle.

Proof: Let $C_m = (v_0, v_1, \ldots, v_{m-1}, v_0)$ be a cycle of smallest size in G_f. Based on C_m, we will consider two consumers whose optimal mechanisms contain as sub-matrices the matrices X, Y from the proof of Claim 16, and hence they cannot be derived from a single mechanism.

The First Consumer: uniform prior over C_m: Consider an information consumer with loss function ℓ_{bin} and prior $p_{v_0} = p_{v_1} = \cdots = p_{v_{m-1}} = 1/m$ and $p_u = 0$ for every $u \notin C_m$. We will construct an optimal mechanism X' for this consumer, and will prove that (in some sense) it is unique. We begin with a labeling algorithm of the vertices in G:

1) Given $C_m = (v_0, v_1, \ldots, v_{m-1}, v_0)$, set $l(v_i) = i$ for $i \in \{0, \ldots, m-1\}$.
2) For s from 1 to $m-1$:
   a) Let V_s be the set of unlabeled vertices that are adjacent to vertices labeled $s-1$.
   b) Let $l(u) = s$ for all $u \in V_s$.
3) Let $l(u) = m-1$ for all remaining vertices u.

Claim 17. *After applying the algorithm, the labels for every two adjacent vertices differ by at most 1 (modulo m).*

Proof: We show that at any stage of the labeling, any two adjacent vertices satisfy the requirement that their labels differ by at most 1 (modulo m).

Note first that this holds for all labeled vertices after Step 1. Consider a vertex $u \in V_s$ (i.e., $l(u) = s$ is set in iteration s), and an adjacent vertex u' that is labeled $l(u') = s'$ prior to or on iteration s. Clearly, if $u' \in V_s \cup V_{s-1}$ then $s' \in \{s-1, s\}$ and the statement holds for (u, u'). Otherwise, we consider two sub-cases. In the first, $l(u') = s' < s-1$, and we are led to a contradiction since u remains unlabeled after iteration $s'+1$ whereas by definition $u \in V_{s'+1}$. In the second sub-case $l(u') = s' > s+1$ (if $s' = s+1$ the claim holds) and hence it must have been that u' was labeled in Step 1, i.e., $u' = v_{s'}$ for $s' \in \{s+1, \ldots, m-1\}$. Following the path of labels which led to the label of u we can get to the vertex v_0 via a path of length s. Noting that this path is disjoint from the length $m - s'$ path $v_{s'} \rightsquigarrow v_0 = v_{s'}, v_{s'+1}, \ldots, v_{m-1}, v_0$, we get that G contains the cycle $v_{s'} \rightsquigarrow v_0 \rightsquigarrow u \rightsquigarrow u'$ that is of length $m - s' + s + 1 < m$, in contradiction to C_m being the smallest cycle in G. To conclude the proof, note that every vertex $u \in G$ adjacent to some $u' \in G$ such that $l(u') \in \{0, 1, \ldots, m-2\}$ has been labeled in iteration $l(u') + 1$ or earlier. Therefore in Step 3, the vertices which are not labeled yet are adjacent only to unlabeled vertices and to vertices with label $m-1$. Labeling the remaining vertices with $m-1$ satisfies the requirement. ■

We now use the graph labels to construct an optimal mechanism X', represented by a matrix of dimensions $|\mathcal{R}_f| \times |\mathcal{R}_f|$. The entries of rows $u \notin C_m$ have no effect on the expected loss of this consumer, as $p_u = 0$. There are, however, restrictions on these rows, as the mechanism X' must be differentially private. We construct X' as follows:

1) For all $u \notin C_m$, set column u of X' to be a column of zeros.
2) For all $u \in C_m$, set row u of X' as in the optimal mechanism X described in the proof of Claim 16 (i.e., as seen in (13)).
3) For all $u \notin C_m$, set row u of X' to be identical to the row corresponding with the vertex identically labeled in C_m.

Clearly, the resulting mechanism is row-stochastic. The privacy constraints also hold: suppose $u, u' \in \mathcal{R}_f$ are query results of neighboring databases. Therefore, they are adjacent in the constraint graph, and their labels differ by at most 1 (modulo m). And so, their matching rows in mechanism X' are either identical or they are the same as rows of two adjacent vertices $v_i, v_j \in C_m$ in mechanism X. Since the construction of rows in X hold to the privacy constraints, so do the rows of X'. In other words, we just showed that mechanism X can be extended to any query f whose constraint graph G_f contains C_m but no smaller cycles.

Notice that only rows of C_m affect the expected loss in X', which is hence identical to that of X. Since any mechanism in this scenario has to satisfy all the restrictions for just the vertices of the cycle C_m, and more, the expected loss for any optimal mechanism in the current scenario is lower bounded with that of X. Hence, we can conclude that X' is optimal for the information consumer, and furthermore, X' restricted to the rows corresponding to C_m is unique.

The Second Consumer: uniform prior over v_0, v_1, v_2: Consider an information consumer with loss function ℓ_{bin} and prior $p_{v_0} = p_{v_1} = p_{v_2} = 1/3$ and $p_u = 0$ for every other $u \in \mathcal{R}_f$. We argue that every optimal mechanism Y' for this consumer has the same structure on rows v_0, v_1, v_2 as mechanism Y in (13). As the impossibility of universal optimality for the case of $m = 3$ was already covered, and we assumed $m > 3$, v_0 and v_2 are not adjacent in G_f. This enables us to label the vertices in such a way: $l(v_0) = 0$, $l(v_2) = 2$ and $l(u) = 1$ for any other vertex in G_f. Again, it is clear that every two adjacent vertices have labels which differ by 1 at most. Similar arguments as the ones presented for the first consumer, show that the first three rows of every optimal mechanism for this consumer (i.e. the rows for v_0, v_1, v_2) have the same structure as the first three rows of mechanism Y in (13).

Assume towards a contradiction that both X' and Y' are derivable from a single mechanism U'. Therefore there exist remappings T, S such that $X' = U'T$ and $Y' = U'S$. Let U be the mechanism U' reduced to only the inputs of the cycle $C_m = \{v_0, v_1, \ldots, v_{m-1}\}$. Reducing U' to U, we get that $X = UT$ and $Y = US$. According to the previous subsection these two mechanisms cannot be derived from a single oblivious mechanism, due to the same arguments in the proof of Claim 11. Thus, we get a contradiction. ■

C. Impossibility of Universal Optimality When $\Delta(G_f) \geq 3$

Finally we focus on acyclic constraint graphs and consider Theorem 15 and its conclusion that for $\alpha > 0.5$ no universally optimal mechanisms exists unless the constraint graph is a simple chain. The proof is omitted and will appear in the full version of this article.

ACKNOWLEDGMENT

Research partially supported by the Israel Science Foundation (grant No. 860/06) and the Frankel Center for Computer Science at Ben-Gurion University.

REFERENCES

[1] A. Blum, K. Ligett, and A. Roth, "A learning theory approach to non-interactive database privacy"., in STOC 2008, pp. 609–618.

[2] I. Dinur and K. Nissim, "Revealing information while preserving privacy", in PODS 2003, pp. 202–210.

[3] C. Dwork, "The differential privacy frontier (extended abstract)", in TCC 2009, pp. 496–502.

[4] C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum, "Differential privacy under continual observation", in STOC 2010, pp. 715–724.

[5] C. Dwork and K. Nissim, "Privacy-preserving datamining on vertically partitioned databases", in CRYPTO 2004, pp. 528–544.

[6] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis", in TCC 2006, pp. 265–284.

[7] C. Dwork, M. Naor, O. Reingold, H. N. Rothblum, and S. P. Vadhan, "On the complexity of differentially private data release: efficient algorithms and hardness results", in STOC 2009, pp. 381–390.

[8] D. Feldman, A. Fiat, H. Kaplan, and K. Nissim, "Private coresets", in STOC 2009, pp. 361–370.

[9] A. Ghosh, T. Roughgarden, and M. Sundararajan, "Universally utility-maximizing privacy mechanisms", in STOC 2009, pp. 351–360.

[10] M. Gupte and M. Sundararajan, "Universally optimal privacy mechanisms for minimax agents", in PODS 2010, pp. 135–146.

[11] D. Kifer and B. R. Lin, "Towards an axiomatization of statistical privacy and utility", in PODS 2010, pp. 147–158.

[12] F. McSherry and K. Talwar, "Mechanism design via differential privacy", in FOCS 2007, pp. 94–103.

[13] K. Nissim, S. Raskhodnikova, and A. Smith, "Smooth sensitivity and sampling in private data analysis", in STOC 2007, pp. 75–84.

[14] A. Roth and T. Roughgarden, "Interactive privacy via the median mechanism", in STOC 2010, pp. 765–774.