On the Geometry of Two-Party Differentially-Private Protocols

Vipul Goyal[*] Ilya Mironov[†] Omkant Pandey[‡] Amit Sahai[§]

Abstract

Differential privacy (DP) is a well-studied notion of privacy, that is generally achieved by randomizing outputs to preserve the privacy of the input records. A central problem in differential privacy is how much accuracy must be lost in order to preserve input privacy?

We study this question in the context of distributed two-party differentially private protocols, where the input is split between two parties. The recent work of McGregor et al. provided several examples of functionalities for which there is an accuracy gap between the client-server setting and the distributed setting. However, many questions remain: does such a gap exist for *any* non-trivial functionality? How large must this gap be? Answering these questions for a large and natural class of functionalities in the two-party setting is the main focus of this work.

Our work obtains general lower bounds on accuracy for differentially private protocols computing any Boolean function. Our bounds are independent of the number of rounds and the communication complexity of the protocol, and hold with respect to computationally unbounded parties. At the heart of our results is a new general geometric technique for obtaining non-trivial accuracy bounds for any Boolean functionality. We obtain the following results:

- We show that for *any* Boolean function, there is a constant accuracy gap between the accuracy that is possible in the client-server setting and the accuracy that is possible in the two-party setting.
- In particular, we show *tight* results on the accuracy that is achievable for the AND and XOR functions in the two-party setting, completely characterizing which accuracies are achievable for any given level of differential privacy.
- Finally, we consider the situation if we relax the privacy requirement to computational differential privacy. We show that to achieve any noticeably better accuracy than what is possible for differentially private two-party protocols, it is essential that one-way functions exist.

[*] Microsoft Research India, Bangalore. Email: vipul@microsoft.com.

[†] Microsoft Research Silicon Valley, Mountain View, CA. Email: mironov@microsoft.com.

[‡] UT Austin. Email: omkant@cs.utexas.edu.

[§] Department of Computer Science, UCLA, Los Angeles, CA. Email: sahai@cs.ucla.edu.

1 Introduction

Differential privacy (DP) is a theoretically sound and well studied notion of privacy (see [Dwo11] for a recent survey). Differential privacy mechanisms work by randomizing the output to preserve the privacy of the input records. A central problem in differential privacy is how much accuracy must be lost in order to preserve input privacy?

The problem of the output being of lower quality appears significantly more difficult in the distributed setting compared to the client-server setting. In the client-server setting, the server can see the entire input in the clear, compute the output correctly and then perturb the output by an appropriate amount to preserve the privacy of the individual entries. However if the input is distributed across several parties, output needs to be computed through an interactive protocol. Throughout the protocol, parties are restricted in how much information their messages should reveal about their input, and this would seem to degrade the quality of the output.

The notion of differential privacy has been studied in the distributed setting, starting with the seminal work of Dwork and Nissim [DN04]. In their work, there are multiple parties each holding a dataset as the input. The parties are interested in mining some information from the entire input vector while preserving the privacy of each individual entry in each dataset. This is a natural setting where the parties may have a legal obligation to protect the privacy of each individual entry (e.g., when the entries consist of medical records). At the same time, the accuracy of the output is very important for the entire computation to be meaningful.

The study of limitations on accuracy of distributed differentially-private protocols was initiated in works of Beimel, Nissim, and Omri [BNO08] for the case of n parties each holding its own input and of McGregor, Mironov, Pitassi, Reingold, Talwar, and Vadhan [MMP$^+$10] for the setting of two parties with n-bit inputs. The latter work considers several natural and constructed functionalities that exhibit a stark gap in accuracy that can be as large as $\Theta(n)$ between client-server and two-party protocols.

However, many questions remain. While we have several examples of functionalities for which there is an accuracy gap between the client-server setting and the distributed setting, does such a gap exist for *any* non-trivial functionality? How large must this gap be? Answering these questions for a large and natural class of functionalities in the two-party setting is the main focus of this work.

Boolean Functionalities. In this work, we focus on protocols that attempt to compute a Boolean function. While much work in differential privacy has focused on computing statistics, we note that computing Boolean functions has long been a motivating goal in differential privacy (e.g., answering the question "Does smoking cause cancer?"). Our goal is to obtain a characterization of which Boolean functions must suffer accuracy loss in the two-party setting, as well as lower bounds on how much accuracy loss is inherently needed. We note that our understanding of Boolean functions has been particularly (and perhaps surprisingly) weak: Before this work, even for computing simple Boolean gates, like AND and XOR, we did not understand whether *any* accuracy loss is essential to the two-party setting.

1.1 Our Results

Before we describe our results, we must define the notion of accuracy that we measure. Since our focus is on functions with Boolean output, there is only one natural choice for accuracy measure: the probability that the output is correct. We note that other metrics considered in the literature do not apply to the Boolean setting.

Now we discuss our setting in more detail. There are two parties Alice and Bob holding inputs x and y respectively and interested in computing a Boolean function $f(x, y)$. The protocol should be such that the differential privacy of each bit of x as well as of y should be preserved[1]. We assume that Alice and Bob follow the protocol as specified, but keep a record of what transpired during the protocol (i.e. they are semi-honest in the cryptographic sense).

For a protocol to achieve accuracy a, it must be the case that for any possible inputs (x, y) to the protocol, the protocol computes the correct output with probability at least a, over the coins of the protocol. Informally speaking, the differential privacy (DP) constraint for Alice states that for any two inputs x_0, x_1 for Alice that differ only in one bit, and for any input y for Bob, the following must hold: For every possible execution of the protocol, the resulting view v of Bob must be such that the probability that v arises on inputs (x_0, y) is within a multiplicative factor of e^ϵ from the probability that v arises on inputs (x_1, y) (see Section 2 for the formal definition of differential privacy). Thus, no matter what Bob sees, he remains uncertain about the value of each bit of Alice's input even if he knows every other bit in her input. Here ϵ is the key privacy parameter. We will denote by λ the value e^ϵ. It is easy to see that in the client-server setting, an accuracy of $\frac{\lambda}{1+\lambda}$ is always possible.

Our work obtains general lower bounds on accuracy. Our bounds are independent of the number of rounds and the communication complexity of the protocol, and hold with respect to computationally unbounded parties. At the heart of our results is a new general geometric technique for obtaining non-trivial accuracy bounds for any Boolean functionality.

General Boolean Functions. Our strategy to obtain results on two-party differentially private protocols for general Boolean functions begins by reducing the problem of obtaining lower bounds for general Boolean functions to specific, simple functions. We first note that Boolean functions where one party's input completely determines the output can of course be computed just as accurately in the two-party setting as in the client-server setting. We call such functions trivial[2]. We then show that the existence of an ϵ-DP protocol with accuracy a for any *non-trivial* function implies the existence of an ϵ-DP protocol with accuracy a for either the AND or XOR functionalities (defined below). Thus, if we can obtain lower bounds on accuracy for AND and XOR, we obtain lower bounds on accuracy for all non-trivial Boolean functions.

Computing an AND gate. The AND functionality is as follows: Alice and Bob each hold a bit denoted by x and y respectively and are interested in computing the AND of the two bits. Given the output (and the protocol transcript), each input bit should remain private. The best known differentially private protocol for this task is the randomized response protocol: each party individually perturbs its input and sends it out. The parties then compute the output based on the two input bits appearing in the protocol transcript. It is easy to see that the output and the protocol transcript still maintain privacy of each individual bit. The randomized response technique gives protocols for AND with accuracy $\frac{\lambda^2}{(1+\lambda)^2}$. The question we consider is: is it optimal?

Towards that end, we show, somewhat surprisingly, that is not the case. We show that a technique we call "shifted randomized response" can achieve higher accuracy than the above randomized response technique. The shifted randomized response technique can achieve an accuracy

[1] Note that stronger notions of privacy are also interesting: for example, where the entire input of each party should be protected, or where symbols larger than bits are to be protected. However, since our focus is on obtaining lower bounds, we use the weaker notion of the privacy stated here, with respect to bits.

[2] This is the same terminology used in works on classifying which Boolean functions have statistically-secure two-party *secure function evaluation* protocols.

of $\frac{\lambda(\lambda^2+\lambda+2)}{(1+\lambda)^3}$. Moreover, we show that this accuracy is *optimal* for AND. We show that any protocol (having any unrestricted number of rounds) can achieve an accuracy of at most $\frac{\lambda(\lambda^2+\lambda+2)}{(1+\lambda)^3}$.

Computing an XOR gate. The XOR functionality is defined analogously to the AND functionality above, except the XOR of the two input bits is to be computed. For the XOR case, the randomized response technique provides an accuracy of $\frac{1+\lambda^2}{(1+\lambda)^2}$. We show that this is, in fact, optimal for XOR.

Combining the results above, we establish the following: *There does not exist any non-trivial Boolean functionality which can be computed with a differential private protocol in the two party setting with accuracy matching that of the client-server setting.* In fact, we obtain a constant separation between the level of accuracy obtainable in the client-server setting and the two-party setting for every non-trivial Boolean functionality, where this constant separation is the best possible in the case of AND and XOR. Our bounds are shown in Figure 1.

Computational Differential Privacy: What assumption is necessary? One option to restore accuracy in a distributed setting is to resort to a relaxed computational notion of differential privacy [MPRV09]. In computational differential privacy (CDP), we relax the privacy condition to require that no *efficient* adversary can predict any bit of the input with probability greater than $\frac{\lambda}{1+\lambda}$, even if the adversary knows all other bits. We ask the question: what computational assumptions are necessary for CDP to enable greater accuracy?

We show that to achieve *any* noticeably greater accuracy with CDP protocols than what is possible with DP protocols, *one-way functions are required*. We show this by presenting a more general result, showing that if one-way functions do not exist, then any CDP protocol must in fact also be a DP protocol.

When discussing CDP protocols, it is important to consider the relationship between CDP protocols and secure computation protocols from cryptography [Yao82, GMW87]. These notions are quite different at a basic syntactic level: for instance, the CDP definition that we consider does not require any two distributions to be computationally distinguishable, whereas this is crucial to generally accepted notions of security for two-party secure computation protocols (both in the input-indistinguishability flavor of definitions as well as the more widely used simulation-based definitions). This is not surprising given the essential philosophical differences between these notions:

- In (computationally) differentially private protocols, “privacy comes first”. We would like to first ensure privacy of each individual input and then with this constraint, would like to compute an accurate output.

- In secure computation protocols, “accuracy comes first”. We would like to release an accurate output to the function we are computing first and then with this constraint, would like to ensure privacy of inputs.

Nevertheless, it is immediate that general secure computation methods do give a way to achieve the same level of accuracy in CDP two-party protocols as in the client-server setting. To achieve this, a secure computation can be used to compute the algorithm that the server would perform on the joint input in the client-server setting. However, general secure computation is essentially equivalent to secure Oblivious Transfer [Kil88, IPS08]. It remains an important open question whether accuracy beyond what is possible with DP protocols may be possible based on an assumption somewhere between one-way functions and the existence of secure Oblivious Transfer protocols.

1.2 Our Techniques.

We present a new general geometric technique for bounding the accuracy of differentially private protocols. At a high level, our technique gives us a method for taking the truth table of a function f, a privacy parameter ϵ, and an accuracy level a, and converting this into a linear program P. We prove that if there does exist an ϵ-DP protocol for computing f with accuracy a, then this linear program must have a solution. By analyzing this LP in the case of specific functions, we can show that no solution exists when a is greater than a bound a^*. This proves that no ϵ-DP protocol can exist with accuracy greater than a^*.

For simplicity, let us focus on protocols for Boolean functions where each party holds a single bit. To obtain our bounds, we first think of every possible "transcript" corresponding to some execution of the protocol. We can associate with each such transcript a 2-by-2 "transcript matrix," whose entries are the probability that this transcript occurs when Alice and Bob start with a particular pair of inputs. Now, each such transcript has an associated output value. If we sum together all the transcript matrices with output value 0, we get a 2-by-2 "protocol matrix," whose entries show the probability that the protocol outputs 0 when Alice and Bob start with a particular pair of inputs.

Now let us consider what constraints we can place on these matrices. Two types of constraints are immediate: (1) the differential privacy conditions on each input linearly constrain each transcript matrix; and (2) the accuracy conditions linearly constrain the protocol matrix. But these constraints alone would not yield any bound better than $\frac{\lambda}{1+\lambda}$, which is achievable in the client-server setting. The key to obtaining better bounds, and our main obstacle, are conditions which capture the constraint that these matrices must actually arise from a *protocol* between two players. We consider a condition that we call *protocol compatibility* that essentially captures the fact that if the two parties' inputs are drawn from independent distributions, then they must remain independent even when conditioned on any particular protocol transcript. This post-execution independence has been useful in other works on differential privacy including the work of McGregor et al. [MMP$^+$10], as well as in works on secure computation such as the work of Kilian [Kil00].

The protocol-compatibility constraint manifests itself as a non-linear (quadratic) constraint on transcript matrices. Note that there can be an enormous (exponential in communication complexity) number of possible transcript matrices, and we do not want to have to consider such a large space of variables. In particular, we do not want our bounds to depend in any way on the communication complexity or the number of rounds in the protocol. We avoid this by proving a key lemma that shows how to optimally combine the linear differential privacy constraints with the non-linear protocol-compatibility constraint to yield a new linear constraint (Lemma 2 in Section 3). This combined linear constraint establishes an upper bound on sums of probabilities from a transcript matrix that combine both the upper and lower bounds from the differential privacy constraints. Because these constraints are linear, they immediately give constraints on the protocol matrix, as well. This gives us our linear program.

We analyze the linear programs that arise specifically for the AND and XOR functionalities, and prove that the linear program is not satisfiable when the accuracy a is higher than a certain value. We prove that these bounds are tight by showing that this accuracy can be achieved for both the AND and XOR functionalities. We stress that our technique is more general, and can be applied to other specific functions to obtain potentially stronger bounds. (As mentioned above, we focus our attention on AND and XOR because every non-trivial Boolean function must contain an embedded AND or XOR function.)

Related Work. In addition to the works mentioned above, several other works have focused on the issue of accuracy and privacy. In the client-server setting (i.e., where only one party owns the entire database), limitations for a wide class of private algorithms were first shown by Dinur and Nissim [DN03], refined and extended by [DMT07, DY08, KRSU10]. Dwork et el. [Dwo06, DMNS06] introduced the concept of differential privacy and provided a general method for achieving non-trivial accuracy based on the sensitivity of the function. The optimality of such mechanisms has since been studied in different models such as answering multiple linear queries [HT10] or producing synthetic datasets by computationally-bounded curators [DNR+09, UV11]. In a surprising result of Ghosh et al. [GRS09], a simple geometric mechanism (a discrete version of the additive Laplacian mechanism) was shown to be universally optimal for releasing a single count query to Bayesian consumers (extended to risk-averse agents by Gupte and Sundararajan [GS10]). Brenner and Nissim [BN10] showed that such universal mechanisms do not exist for other types of queries.

In the secure function evaluation model against computationally unbounded semi-honest adversaries, characterization of deterministic Boolean functionalities was completed by Chor and Kushilevitz [CK91], and for randomized functionalities by [Kil00]. These results establish the "all or nothing" nature of two-party computation under information-theoretic reductions. A related question of characterizing complete *deterministic* functionalities in the computational setting was considered by Harnik et al. [HNRR06]. Complete classification of randomized functionalities in the computational setting remains an important research problem.

2 Notation and Definitions

Standard notation. We use symbols $\neg, \vee, \wedge$, and $\oplus$ to denote the standard Boolean operations: NOT, OR, AND, and XOR respectively. The set of natural numbers is denoted by $\mathbb{N}$; for $n \in \mathbb{N}$, we write by $[n]$ as shorthand for the set $\{1, 2, \ldots, n\}$. The Hamming distance between two strings $x, y \in \{0,1\}^n$ is defined as: $|x - y|_h = |\{i \in [n] \colon x_i \neq y_i\}|$, where x_i, y_i denote the ith bit of x, y respectively. We denote by e, the base of the natural logarithm.

We now recall the definition of ϵ-differential-privacy [Dwo06] and (ϵ, δ)-differential privacy [DKM+06].

Definition 1 (ϵ-Differential-Privacy) *A randomized function $M \colon \{0,1\}^n \mapsto \mathcal{R}$, with a finite range $\mathcal{R}$, is said to be an ϵ-differentially-private (ϵ-DP) mechanism for $\epsilon \geq 0$ if for every $(x, y) \in \{0,1\}^n \times \{0,1\}^n$ satisfying $|x - x'|_h = 1$ and every subset $S \subset \mathcal{R}$ we have that over the randomness of M:*

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

Definition 2 ((ϵ, δ)-Differential-Privacy) *A randomized function $M \colon \{0,1\}^n \mapsto \mathcal{R}$, with a finite range $\mathcal{R}$, is said to be a (ϵ, δ)-differentially-private mechanism for $\epsilon, \delta \geq 0$ if for every $(x, y) \in \{0,1\}^n \times \{0,1\}^n$ satisfying $|x - x'|_h = 1$ and every subset $S \subset \mathcal{R}$ we have that over the randomness of M:*

$$\Pr[M(x) \in S] \leq e^{\epsilon} \times \Pr[M(x') \in S] + \delta.$$

We next recall the definition of computational differential privacy which captures differentially privacy for polynomial time tests. We work with the weakest definition, namely ϵ-IND-CDP [MPRV09]. In the following, k denotes the security parameter, implicitly available to all algorithms; algorithms are assumed to run in time polynomial in k unless stated otherwise.

Definition 3 (ϵ-IND-CDP Privacy) *An ensemble $\{M_\kappa\}_{k \in \mathbb{N}}$ of randomized functions $M_k \colon \{0,1\}^n \mapsto \mathcal{R}_k$ provides ϵ-IND-CDP if there exists a negligible function $\mathrm{negl}(\cdot)$ such that for every probabilistic polynomial time distinguisher A, for every polynomial $p(\cdot)$, for any adjacent strings $x, x' \in \{0,1\}^n$*

(i.e., $|x - x'|_h = 1$), for every sufficiently large $k \in \mathbb{N}$, and for every advice string z_k of size at most $p(k)$, it holds that

$$\Pr\left[A_k(M_k(x)) = 1\right] \leq e^{\epsilon} \times \Pr\left[A_k(M_k(x')) = 1\right] + \mathrm{negl}(\kappa),$$

where we write $A_k(x)$ for $A(1^k, z_k, x)$ and the probability is taken over the randomness of mechanism M_k and the distinguisher A.

Interactive Setting. Let $\pi := \langle A, B \rangle$ be a two-party protocol. Define $\textsc{view}^A_\pi(x, y)$ to be the random-variable which, in a random-execution of π with inputs x, y for A, B respectively, consists of (x, R_A, trans), where R_A is the randomness used by A and trans is the sequence of messages exchanged between the parties in the sampled execution. For each x, $\textsc{view}^A_\pi(x, y)$ is a mechanism over the y's. Define $\textsc{view}^B_\pi(x, y)$ analogously. When dealing with the computational notion, we consider the family of protocols $\{\pi_k\}_{k \in \mathbb{N}}$ and denote the view of A (resp., B) by $\textsc{view}^A_\pi(k, x, y)$ (resp., $\textsc{view}^B_\pi(k, x, y)$).

Definition 4 (Two Party Differentially Privacy) *We say that a protocol $\pi := \langle A, B \rangle$ is ϵ-DP (resp., (ϵ, δ)-DP) if the mechanism $\textsc{view}^A_\pi(x, y)$ is ϵ-DP (resp., (ϵ, δ)-DP) for all values of x and the same holds for $\textsc{view}^B_\pi(x, y)$. Analogously, a family of protocols $\{\pi_k\}_{k \in \mathbb{N}}$ is ϵ-IND-CDP if the mechanism $\textsc{view}^A_\pi(k, x, y)$ is ϵ-IND-CDP for all values of x and every sufficiently large k, and the same holds for $\textsc{view}^B_\pi(k, x, y)$.*

Finally, our measure of accuracy for Boolean functions simply looks at how often a randomized mechanism outputs the correct output bit in the worst case.

Definition 5 (Accuracy) *The accuracy of (randomized) mechanism $M\colon \{0,1\}^n \mapsto \{0,1\}$ with respect to a Boolean function $f\colon \{0,1\}^n \mapsto \{0,1\}$ is defined as: $\mathsf{Acc_f}(M) = \min_x\{\Pr[M(x) = f(x)]\}$, where the probability is taken over the randomness of M.*

The accuracy of a two party protocol $\pi := \langle A, B \rangle$ w.r.t. to $f\colon \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$ is defined as the accuracy of the mechanism $\textsc{out}_\pi\colon \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$ which returns the (official) output of the protocol in a randomly sampled execution of π. The accuracy for a family of protocols $\{\pi_k\}_{k \in \mathbb{N}}$ is defined analogously for each k.

3 Geometric Analysis of Two-party Differential Privacy for Boolean Functions

We first show that AND and XOR gates are *embedded on adjacent inputs* into any non-trivial Boolean two-party functionality, i.e., any functionality whose output is not fully determined by one side's input (Section 3.1). Therefore, it will be sufficient to analyze AND/XOR gates. A similar claim also appears in [CKL03] but does not guarantee the *adjacency* of inputs that embed AND/XOR gates. Adjacency is crucial in our case, since otherwise we cannot conclude that the protocol for AND/XOR have the same privacy parameter ϵ.

We then formulate necessary conditions for existence of a differentially-private two-party protocol implementing a randomized two-party Boolean functionality (Section 3.2), and use these conditions towards tight analysis of accuracy of AND and XOR gates achievable via differentially-private protocols (Sections 3.3 and 3.4).

The main results of the section are plotted numerically in Figure 1.

3.1 Embedded XOR and AND functionalities

We show that every non-trivial Boolean function $f(x,y)$ defined over $x,y \in \{0,1\}^n$ for $n \in \mathbb{N}$, embeds either an AND or an XOR function on two *adjacent* inputs. Function f is trivial if it does not depend on the inputs of both parties.

Definition 6 *A function $f\colon \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$ is said to be* trivial *if there exists a function $g\colon \{0,1\}^n \mapsto \{0,1\}$ such that*

$$\textit{for all } x,y \in \{0,1\}^n \quad f(x,y) = g(x)$$

or

$$\textit{for all } x,y \in \{0,1\}^n \quad f(x,y) = g(y).$$

If f is not a trivial function, we say that f is non-trivial.

Definition 7 *A function $f\colon \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$* embeds the XOR function on two adjacent inputs *if there exists a bit b, and inputs $x,x',y,y' \in \{0,1\}^n$ such that x and y are adjacent to x' and y' respectively (i.e. $|x-x'|_h = |y-y'|_h = 1$) and it holds that:*

$$\begin{aligned} f(x,y) =& f(x',y') = b, \\ f(x,y') =& f(x',y) = 1-b. \end{aligned}$$

We say that f embeds the AND function on two adjacent inputs *if*

$$\begin{aligned} &f(x,y) = b, \\ &f(x,y') = f(x',y) = f(x',y') = 1-b. \end{aligned}$$

Lemma 1 *Every non-trivial function $f\colon \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$ embeds either the XOR or the AND function on two adjacent inputs.*

Proof: Assume the opposite: the function f is non-trivial but does not embed the XOR or the AND on adjacent inputs.

Let $N = 2^n$ and $(g_1,\ldots,g_N)$ be the Gray encoding of all binary vectors in $\{0,1\}^n$. Recall that a Gray encoding is a permutation over $\{0,1\}^n$ with the property that for every $i \in [N]$, g_i and g_{i+1} are adjacent vectors: $|g_i - g_{i+1}|_h = 1$. Let T denote the truth table of f arranged in the Gray encoding sequence, i.e., T is an $N \times N$ matrix such that $T(i,j) = f(g_i, g_j)$.

We colors cells of T black if they hold a zero, and white otherwise. We say that a row (resp., column) of T is *monochromatic* if all entries of that row (resp., column), are of the same color.

It is easy to verify that f is trivial if and only if either all rows or all columns of T are simultaneously monochromatic. Therefore, if f is non-trivial then there must exist a row that is not monochromatic.

We claim that two adjacent rows of T are either identical or the negatives of each other. Indeed, consider two adjacent rows i and $i+1$ and label the column j with '=' if $T(i,j) = T(i+1,j)$ and '$\neq$' if $T(i,j) \neq T(i+1,j)$. If there are two columns anywhere labeled with different symbols, there must be two *adjacent* columns j and $j+1$ labeled with '=' and '$\neq$'. Without loss of generality, $T(i,j) = T(i+1,j)$ and $T(i,j+1) \neq T(i+1,j+1)$. It is immediate that f embeds the AND function on inputs $g_i, g_{i+1}, g_j, g_{j+1}$. Therefore, all columns must all be labeled with either '=' or '$\neq$', which means that the adjacent rows must be either be identical or the negatives (inverses) of each other.

Moreover, if the row i is not monochromatic, then its adjacent rows $i+1$ and $i-1$ must be identical to it. As we have shown, the alternative is for the adjacent row to be the negative of the ith row. Considering two adjacent entries of the ith row that are of different colors, we find that in that case f must embed the XOR function.

We just proved that if there is a single non-monochromatic row, then its adjacent rows, and by induction, all rows of the matrix are equal to it. Since T contains a non-monochromatic row, it means that all its rows are identical, and all columns are monochromatic. This is a contradiction since f is non-trivial. ▮

3.2 Differential privacy and protocol compatibility

We begin by introducing several definitions pertaining to properties of matrices that describe joint distributions of protocol outcomes as a function of two inputs. For compactness we will use $\lambda = e^{\epsilon}$ without stating it explicitly throughout the section.

Definition 8 (ϵ-DP matrix) *A $2^n \times 2^n$ matrix P indexed by strings $x, y \in \{0,1\}^n$ is* ϵ-DP *if its elements satisfy the following conditions for all adjacent pairs $x, x' \in \{0,1\}^n$ and $y, y' \in \{0,1\}^n$:*

$$p_{xy} \leq \lambda \cdot p_{xy'},$$
$$p_{xy} \leq \lambda \cdot p_{x'y},$$

Definition 9 (Protocol-compatible matrix) *A $2^n \times 2^n$ matrix P is* protocol compatible *if for all $x_1, x_2, y_1, y_2 \in \{0,1\}^n$ it holds that*

$$p_{x_1y_1} \cdot p_{x_2y_2} = p_{x_1y_2} \cdot p_{x_2y_1}$$

The next two definitions extend the concepts of differential privacy and protocol compatibility to two-party Boolean functionalities. By convention, we say that a $2^n \times 2^n$ matrix P represents a randomized Boolean functionality f of two inputs if $p_{xy} = \Pr[f(x,y) = 0]$ for all $x, y \in \{0,1\}^n$.

Definition 10 (ϵ-DP functionality) *We call a $2^n \times 2^n$ matrix P an* ϵ-DP functionality *if both P and $\mathbf{1} - P$ are* ϵ-DP *matrices, where $\mathbf{1}$ is the all-ones matrix.*

Definition 11 (Protocol-compatible ϵ-DP functionality) *A $2^n \times 2^n$ matrix P is* protocol-compatible ϵ-DP functionality *if both matrices P and $\mathbf{1} - P$ can be expressed as sums of protocol-compatible* ϵ-DP *matrices, where $\mathbf{1}$ is the all-ones matrix.*

The following theorem establishes necessary conditions for existence of a differentially-private two-party protocol for computing a randomized predicate of two n-bit inputs.

Theorem 1 *Let π be a randomized ϵ-DP two-party protocol defined over $x, y \in \{0,1\}^n$ and $\pi(x,y)$ be the Boolean output of the protocol. Let P be a matrix of probabilities $p_{xy} = \Pr[\pi(x,y) = 0]$. Then P is a protocol-compatible* ϵ-DP *functionality.*

Proof: We start by showing that P can be expressed as sums of protocol-compatible ϵ-DP matrices. The proof for $\mathbf{1} - P$ is analogous.

Let T_0 be the set of all transcripts τ for which the protocol output is 0, i.e. $\pi(x,y) = 0$. For a fixed x, y, let $\tau \leftarrow \pi(x,y)$ denote that event that in a random execution of π with inputs (x,y), the transcript is τ. Let P_τ be a $2^n \times 2^n$ matrix indexed by n-bit strings such that $P_\tau(x,y) = \Pr[\tau \leftarrow \pi(x,y)] = p_{\tau,xy}$ (say). Then,

$$p_{xy} = \Pr[\pi(x,y) = 0] = \sum_{\tau \in T_0} \Pr[\tau \leftarrow \pi(x,y)] = \sum_{\tau \in T_0} p_{\tau,xy}.$$

Therefore, it holds that $P = \sum_{\tau \in T_0} P_\tau$. It is easy to verify that the matrices P_τ are ϵ-DP matrices. To complete the proof, we now show that each P_τ is protocol-compatible (following [Kil00]).

Let X and Y be independently and uniformly distributed random variables taking values in $\{0,1\}^n$. Then, using Bayes' rule we see that for any two strings x, y, $p_{\tau,xy} = \Pr[X = x, Y = y | \tau \leftarrow \pi(X,Y)] \cdot \Pr[\tau \leftarrow \pi(X,Y)] / \Pr[X = x, Y = y]$. It is well known in communication complexity (e.g., see [MMP$^+$10]) that for any two-party protocol π, if the inputs X and Y are independent before the execution, then for any transcript τ of the protocol, X and Y remain independent when conditioned on the transcript being τ. That is, $\Pr[X = x, Y = y | \tau \leftarrow \pi(X,Y)] = \Pr[X = x | \tau \leftarrow \pi(X,Y)] \cdot \Pr[Y = y | \tau \leftarrow \pi(X,Y)]$. Using this with our previous relation, we see that

$$p_{\tau,xy} = p_{x,\tau} \cdot p_{y,\tau} \cdot p_\tau \cdot 2^{2n},$$

where $p_{x,\tau} = \Pr[X = x | \tau \leftarrow \pi(X,Y)]$; $p_{y,\tau}$ and p_τ are defined analogously. It then follows that for any distinct x_1, y_1, x_2, y_2,

$$p_{\tau,x_1y_1} \cdot p_{\tau,x_2,y_2} = p_{x_1,\tau} p_{x_2,\tau} p_{y_1,\tau} p_{y_2,\tau} \cdot p_\tau^2 \cdot 2^{4n} = p_{\tau,x_1y_2} \cdot p_{\tau,x_2,y_1}.$$

This completes the proof for protocol-compatibility, and hence the theorem. ■

The following lemma plays a critical role in our analysis, as it replaces a per-transcript quadratic constraint imposed by the protocol-compatibility condition with a system of linear inequalities.

Lemma 2 *If P is protocol-compatible ϵ-DP functionality, then for all adjacent pairs $x, x' \in \{0,1\}^n$ and $y, y' \in \{0,1\}^n$*

$$\begin{aligned} p_{xy'} + p_{x'y} &\le p_{xy}/\lambda + p_{x'y'} \cdot \lambda, \\ p_{xy'} + p_{x'y} &\le p_{xy} \cdot \lambda + p_{x'y'}/\lambda, \end{aligned}$$

and

$$\begin{aligned} p_{xy} + p_{x'y'} &\le p_{xy'}/\lambda + p_{x'y} \cdot \lambda, \\ p_{xy} + p_{x'y'} &\le p_{xy'} \cdot \lambda + p_{x'y}/\lambda. \end{aligned}$$

Proof: We first verify the statement for protocol-compatible ϵ-DP matrices Q. Indeed, by the ϵ-DP condition $q_{xy'}, q_{x'y} \in [q_{xy}/\lambda, q_{x'y'} \cdot \lambda]$ and by protocol-compatible $q_{xy'} \cdot q_{x'y} = q_{xy} \cdot q_{x'y'}$. If the product of two reals is fixed, their sum is maximized when they are most apart, which corresponds exactly to the endpoints of the feasible interval for $q_{xy'}$, $q_{x'y}$. To formalize this, we observe that by simple algebra, the condition

$$q_{xy}/\lambda \le q_{xy'} \le q_{x'y'} \cdot \lambda$$

is equivalent to the quadratic inequality

$$q_{xy'}^2 - (q_{xy}/\lambda + q_{x'y'} \cdot \lambda) q_{xy'} + q_{xy} \cdot q_{x'y'} \le 0,$$

since all probabilities must be non-negative. Rewriting this inequality and dividing by $q_{xy'}$, and using the fact that $q_{x'y} = q_{xy} \cdot q_{x'y'}/q_{xy'}$, we obtain the desired bound:

$$q_{xy'} + q_{x'y} \leq q_{xy}/\lambda + q_{x'y'} \cdot \lambda.$$

Moreover, the bound is linear in all q_{xy}, $q_{xy'}$, $q_{x'y}$, $q_{x'y'}$ and holds for all protocol-compatible ϵ-DP matrices. Therefore, it would also hold for the sum of these matrices, and thus for protocol-compatible ϵ-DP functionalities. The other bounds follow similarly and this completes the proof. ∎

Lastly, we introduce the following definition that relaxes the notion of the protocol-compatible ϵ-DP functionality to allow for a (typically small or negligible) fraction of non-private transcripts.

Definition 12 (Protocol-compatible ϵ-DP δ-close functionality) *A $2^n \times 2^n$ matrix P is* protocol-compatible ϵ-DP functionality *if both matrices P and $\mathbf{1} - P - \Delta$ can be expressed as sums of protocol-compatible ϵ-*DP *matrices, where $\mathbf{1}$ is the all-ones matrix and all entries of Δ are between 0 and δ.*

An analogue of Theorem 1 exists for (ϵ, δ)-functionalities defined over *binary* inputs:

Theorem 2 *Let π be a randomized (ϵ, δ)-DP two-party protocol defined over $x, y \in \{0, 1\}$ and $\pi(x, y)$ be the Boolean output of the protocol. Let P be a matrix of probabilities $p_{xy} = \Pr[\pi(x, y) = 0]$. Then P is a protocol-compatible $(\epsilon + \sqrt{\delta})$-DP $O(\sqrt{\delta})$-close functionality.*

Proof: In the notation of the previous theorem, define the set of "bad" transcripts B as

$$B = \{\tau \colon \exists \text{ adjacent } x, x', y, y' \in \{0, 1\}, \text{s.t. } P_\tau(x, y) > e^{\epsilon + \sqrt{\delta}} P_\tau(x', y')\}.$$

We claim that for all $x, y \in \{0, 1\}$, the probability that $\Pr[\tau \in B \colon \tau \leftarrow \pi(x, y)] < O(\sqrt{\delta})$. Applying Theorem 1, it is sufficient to prove the claim.

For any two pairs of adjacent inputs x, x', y, y' define

$$B_{x,x',y,y'} = \{\tau \colon P_\tau(x, y) > e^{\epsilon + \sqrt{\delta}} P_\tau(x', y')\}. \quad (1)$$

The probability of seeing a transcript from $B_{x,x',y,y'}$ on input (x, y) is less than $O(\sqrt{\delta})$, since by the guarantee of (ϵ, δ)-DP and (1):

$$\Pr[\tau \in B_{x,x',y,y'} \colon \tau \leftarrow \pi(x, y)\} \leq e^{\epsilon} \Pr[\tau \in B_{x,x',y,y'} \colon \tau \leftarrow \pi(x', y')\} + \delta < \\ < e^{-\sqrt{\delta}} \Pr[\tau \in B_{x,x',y,y'} \colon \tau \leftarrow \pi(x, y)\} + \delta,$$

from which a $O(\sqrt{(\delta)})$ bound on $\Pr[\tau \in B_{x,x',y,y'} \colon \tau \leftarrow \pi(x, y)\}$ follows immediately.

Applying the (ϵ, δ)-DP condition again, we find that $\Pr[\tau \in B_{x,x',y,y'} \colon \pi(x'', y'')]$ for all $x'', y'' \in \{0, 1\}$ is bounded by $e^{\epsilon} O(\sqrt{\delta})$. Since the event B is the union of all events $B_{x,x',y,y'}$, summing over all pairs of adjacent inputs and assuming that ϵ is constant, we complete the proof. ∎

The next two sections apply Theorem 1 to tight analysis of accuracy of differentially private protocols for computing two Boolean functionalities of two bit inputs: AND and XOR.

3.3 Analysis of the AND functionality

We first define accuracy of a Boolean functionality for computing the AND of two bit inputs specified as a 2×2 matrix of probabilities. Recall that by convention, the matrix P consists of elements p_{xy} signifying the probability of obtaining output 0 on inputs (x, y).

Definition 13 (AND-Accuracy) *Define AND-accuracy of a 2×2 matrix $\left(\begin{smallmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{smallmatrix}\right)$ as*

$$\mathsf{AND\text{-}Acc}(\left(\begin{smallmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{smallmatrix}\right)) = \min(p_{00}, p_{01}, p_{10}, 1 - p_{11}).$$

Note that this notion is identical to the accuracy defined in section 2. We prove the following theorem establishing the maximal accuracies achievable by protocol-compatible and arbitrary ϵ-DP functionalities and, in particular, showing that there is a gap between the two quantities.

Theorem 3 *For any $\lambda \geq 1$ and a 2×2 matrix M we have the following:*

1. *If M is a ϵ-DP functionality, then $\mathsf{AND\text{-}Acc}(M) \leq \frac{\lambda}{1+\lambda}$, where $\lambda = e^\epsilon$.*
2. *If M is a ϵ-DP protocol-compatible functionality, then $\mathsf{AND\text{-}Acc}(M) \leq \frac{\lambda(\lambda^2+\lambda+2)}{(1+\lambda)^3}$.*

In both cases the equality can be achieved.

Proof: Let $M = \left(\begin{smallmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{smallmatrix}\right)$ and $a = \mathsf{AND\text{-}Acc}(M)$.

Claim 1. The accuracy condition implies $p_{01} \geq a$ and $1 - p_{11} \geq a$. On the other hand, by the ϵ-DP constraint $p_{01} \leq p_{11} \cdot \lambda$. Put together we have $a/\lambda \leq p_{11} \leq 1 - a$, which implies $a \leq \lambda/(1 + \lambda)$.

The following matrix is indeed a ϵ-DP functionality with accuracy $\lambda/(1 + \lambda)$:

$$M = \begin{pmatrix} \lambda/(1+\lambda) & \lambda/(1+\lambda) \\ \lambda/(1+\lambda) & 1/(1+\lambda) \end{pmatrix}.$$

Claim 2. The following conditions relate the probabilities p_{00}, p_{01}, p_{10}, p_{11} to each other and to the accuracy parameter a:

$$\begin{aligned} p_{11} &\leq 1 - a \\ p_{01} + p_{10} &\geq 2 \cdot a \\ p_{01} + p_{10} &\leq p_{00}/\lambda + p_{11} \cdot \lambda \\ (1 - p_{01}) + (1 - p_{10}) &\leq (1 - p_{00}) \cdot \lambda + (1 - p_{11})/\lambda. \end{aligned}$$

The first two inequalities are implied by the accuracy requirement, the last two by applying Lemma 2 to M and $\left(\begin{smallmatrix} 1-p_{00} & 1-p_{01} \\ 1-p_{10} & 1-p_{11} \end{smallmatrix}\right)$.

By introducing a new variable $q = p_{01} + p_{10}$ and rewriting the above inequalities, we have

$$q \geq 2 \cdot a \tag{2}$$

$$q \leq p_{00}/\lambda + p_{11} \cdot \lambda \tag{3}$$

$$q \geq 2 - (\lambda + 1/\lambda) + p_{00} \cdot \lambda + p_{11}/\lambda. \tag{4}$$

Consider the intersection of two lines bounding the half-planes (2) and (4), where q and p_{00} are considered as free variables and λ, a, and p_{11} are parameters. It is easy to verify that the lines intersect at the point (p_{00}^*, q^*), where

$$p_{00}^* = 1 + 1/\lambda^2 - 2/\lambda + 2a/\lambda - p_{11}/\lambda^2 \quad \text{and} \quad q^* = 2 \cdot a.$$

The following lemma argues that (p_{00}^*, q^*) satisfies (3):

Lemma 3 *Let p_{00}^* be defined as above. Then the following holds:*

$$2 \cdot a \leq p_{00}^*/\lambda + p_{11} \cdot \lambda.$$

Proof [Lemma 3]**:** Towards a contradiction, assume that

$$2 \cdot a > p_{00}^*/\lambda + p_{11} \cdot \lambda. \tag{5}$$

Consider two cases:

Case $p_{00} \leq p_{00}^*$. Then

$$q \overset{(3)}{\leq} p_{00}/\lambda + p_{11} \cdot \lambda \leq p_{00}^*/\lambda + p_{11} \cdot \lambda \overset{(5)}{<} 2 \cdot a,$$

contradicting (2).

Case $p_{00} > p_{00}^*$. Then

$$(p_{00} - p_{00}^*)/\lambda + 2 \cdot a \overset{(5)}{>} p_{00}/\lambda + p_{11} \cdot \lambda \overset{(3)}{\geq} q \overset{(4)}{\geq} 2 - (\lambda + 1/\lambda) + p_{00} \cdot \lambda + p_{11}/\lambda =$$
$$= (p_{00} - p_{00}^*) \cdot \lambda + p_{00}^* \cdot \lambda + 2 - (\lambda + 1/\lambda) + p_{11}/\lambda \overset{\text{def of } p_{00}^*}{=} (p_{00} - p_{00}^*) \cdot \lambda + 2 \cdot a,$$

which is a contradiction since $\lambda \geq 1$ and $p_{00} > p_{00}^*$. ■ [Lemma 3]

Finally, by substituting the value of p_{00}^* into the statement of Lemma 3 and using that $p_{11} \leq 1 - a$, we have

$$2 \cdot a \leq (1 + 1/\lambda^2 - 2/\lambda + 2a/\lambda - p_{11}/\lambda^2)/\lambda + p_{11} \cdot \lambda \leq \lambda + 1/\lambda - 2/\lambda^2 + a \cdot (2/\lambda^2 - \lambda + 1/\lambda^3),$$

from which after collecting like terms and simplifying, the claim $\mathsf{AND\text{-}Acc}(M) = a \leq \lambda(\lambda^2 + \lambda + 2)/(1 + \lambda)^3$ follows.

The protocol with optimal accuracy is shown in Figure 2 thus proving tightness of the bound. ■

We remark that Claim 2 of Theorem 3 also applies to δ-close ϵ-DP protocol compatible functionalities, with the upper bound on the accuracy increasing by $(2 + \lambda + 1/\lambda)\delta = O(\delta)$. The proof changes in its application of Lemma 2 to $\left(\begin{smallmatrix} 1-p_{00} & 1-p_{01} \\ 1-p_{10} & 1-p_{11} \end{smallmatrix}\right)$ that becomes instead $\left(\begin{smallmatrix} 1-p_{00}-\delta_{00} & 1-p_{01}-\delta_{01} \\ 1-p_{10}-\delta_{10} & 1-p_{11}-\delta_{11} \end{smallmatrix}\right)$, where $\delta_{00}, \delta_{01}, \delta_{10}, \delta_{11} \in [0, \delta]$. It is easy to verify that changes in the inequality (4) can be absorbed by reducing the value of a by $(2 + \lambda + 1/\lambda)\delta = O(\delta)$.

Maximal accuracies attained by ϵ-DP functionalities and protocol-compatible ϵ-DP functionalities are shown in Figure 1.

3.4 Analysis of the XOR functionality

Definition 14 (XOR-Accuracy) *Define XOR-accuracy of a 2×2 matrix $\left(\begin{smallmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{smallmatrix}\right)$ as*

$$\mathsf{XOR\text{-}Acc}(\left(\begin{smallmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{smallmatrix}\right)) = \min(p_{00}, 1 - p_{01}, 1 - p_{10}, p_{11}).$$

Note that this notion is identical to the accuracy defined in section 2. The following theorem bounds XOR-accuracy of DP functionalities and protocol-compatible DP functionalities.

Theorem 4 *For any $\lambda \geq 1$ and a 2×2 matrix M we have the following:*

1. *If M is a ϵ-DP functionality, then* $\mathsf{XOR\text{-}Acc}(M) \leq \frac{\lambda}{1+\lambda}$.
2. *If M is a ϵ-DP protocol-compatible functionality, then* $\mathsf{XOR\text{-}Acc}(M) \leq \frac{1+\lambda^2}{(1+\lambda)^2}$.

In both cases the equality can be achieved.

Proof: Let $M = \left(\begin{smallmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{smallmatrix}\right)$ and $a = \mathsf{AND\text{-}Acc}(M)$.

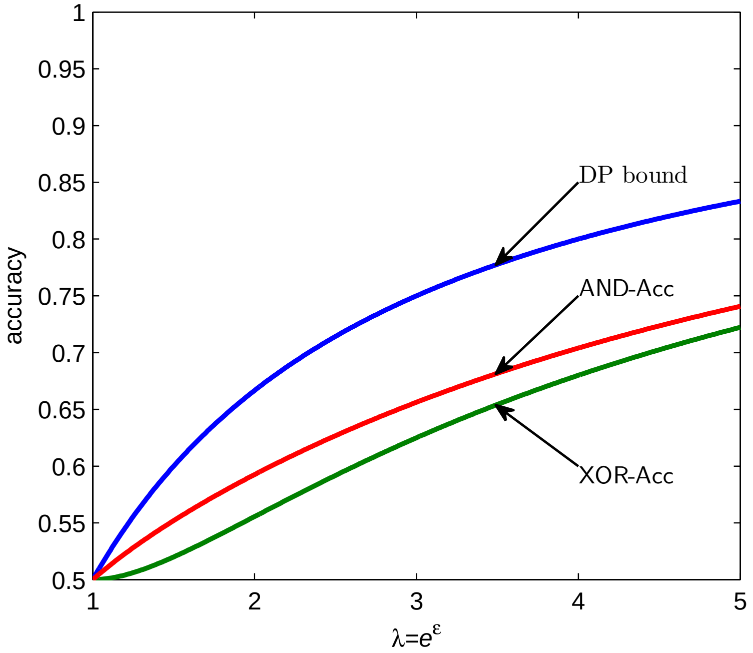


Figure 1: Bounds on accuracy: for arbitrary Boolean functionalities (DP bound); for protocol compatible ϵ-DP AND and XOR. Every ϵ-DP protocol for any non-trivial Boolean functionality must be subject to either the AND or XOR bound.

Claim 1. By the accuracy condition $p_{00} \geq a$ and $1 - p_{01} \geq a$. On the other hand, by the ϵ-DP constraint $p_{00} \leq p_{01} \cdot \lambda$. Put together we have $1 - a \geq p_{01} \geq p_{00}/\lambda \geq a/\lambda$, which implies $a \leq \lambda/(1+\lambda)$.

The following matrix is indeed a ϵ-DP functionality with accuracy $\lambda/(1+\lambda)$:

$$M = \begin{pmatrix} \lambda/(1+\lambda) & 1/(1+\lambda) \\ 1/(1+\lambda) & \lambda/(1+\lambda) \end{pmatrix}.$$

Claim 2. Lemma 2 gives the following bounds on the entries of ϵ-DP protocol-compatible matrices:

$$p_{00} + p_{11} \leq p_{10}/\lambda + p_{01} \cdot \lambda,$$
$$p_{00} + p_{11} \leq p_{10} \cdot \lambda + p_{01}/\lambda.$$

Summing the inequalities and dividing by two, we have

$$p_{00} + p_{11} \leq \frac{\lambda^2 + 1}{2\lambda}(p_{10} + p_{01}). \tag{6}$$

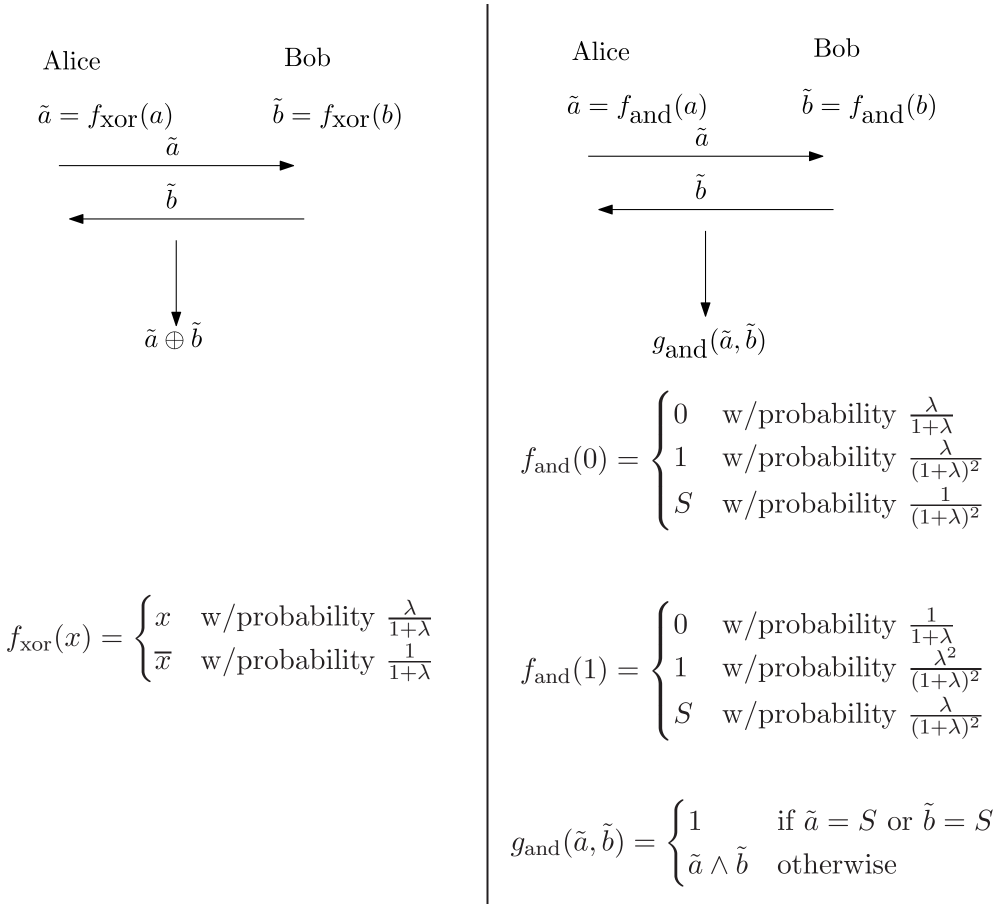
Observe that $\mathsf{XOR\text{-}Acc}(M) = \min(p_{00}, 1 - p_{01}, 1 - p_{10}, p_{11}) \leq \min(\frac{p_{00}+p_{11}}{2}, 1 - \frac{p_{10}+p_{01}}{2})$. Denote $\frac{p_{00}+p_{11}}{2}$ by x and $\frac{p_{10}+p_{01}}{2}$ by y, and write

$$\mathsf{XOR\text{-}Acc}(M) = \min(x, 1-y) \overset{(6)}{\leq} \min(x, 1 - \frac{2\lambda}{1+\lambda^2}x),$$

which attains its maximal value when $x = 1 - \frac{2\lambda}{1+\lambda^2}x$. Solving this for x and substituting in the above expression, we prove that

$$\mathsf{XOR\text{-}Acc}(M) \leq \frac{1+\lambda^2}{(1+\lambda)^2}.$$

This value of accuracy for computing the XOR functionality is achieved by the randomized response protocol, Figure 2. ∎



Alice
Bob
$\tilde{a} = f_{\text{xor}}(a)$
$\tilde{b} = f_{\text{xor}}(b)$
$\tilde{a}$
$\tilde{b}$
$\tilde{a} \oplus \tilde{b}$
Alice
Bob
$\tilde{a} = f_{\text{and}}(a)$
$\tilde{b} = f_{\text{and}}(b)$
$\tilde{a}$
$\tilde{b}$
$g_{\text{and}}(\tilde{a}, \tilde{b})$

$$f_{\text{xor}}(x) = \begin{cases} x & \text{w/probability } \frac{\lambda}{1+\lambda} \\ \overline{x} & \text{w/probability } \frac{1}{1+\lambda} \end{cases}$$

$$f_{\text{and}}(0) = \begin{cases} 0 & \text{w/probability } \frac{\lambda}{1+\lambda} \\ 1 & \text{w/probability } \frac{\lambda}{(1+\lambda)^2} \\ S & \text{w/probability } \frac{1}{(1+\lambda)^2} \end{cases}$$

$$f_{\text{and}}(1) = \begin{cases} 0 & \text{w/probability } \frac{1}{1+\lambda} \\ 1 & \text{w/probability } \frac{\lambda^2}{(1+\lambda)^2} \\ S & \text{w/probability } \frac{\lambda}{(1+\lambda)^2} \end{cases}$$

$$g_{\text{and}}(\tilde{a}, \tilde{b}) = \begin{cases} 1 & \text{if } \tilde{a} = S \text{ or } \tilde{b} = S \\ \tilde{a} \wedge \tilde{b} & \text{otherwise} \end{cases}$$

Figure 2: Optimal XOR and AND protocols.

4 One-Way Functions from CDP

In this paper, we show that one-way functions are implied by the existence of a family of *computationally* differentially private (CDP) two-party protocols that achieve better accuracy than the bounds proven for DP two-party protocols proven in the previous section. We show this by presenting a more general result: we show that if one-way functions *do not exist*, then the existence of a family of CDP protocols imply the existence of DP protocols with only negligible loss in accuracy and privacy.

Definition 15 *An infinite family of two-party protocols* $\Pi = \{\pi_k\}$ *is defined to be an infinite* family of $(\epsilon, \delta = \text{negligible})$-DP protocols achieving accuracy a for a functionality F *if for every constant*

$c > 0$, there exists an infinite sequence of $\pi_k \in \Pi$ such that each π_k is an $(\epsilon + k^{-c}, \delta = k^{-c})$-DP protocol with accuracy $a - k^{-c}$ for functionality F.

Our main theorem is the following:

Theorem 5 *Suppose that one-way functions do not exist. Then given any infinite family Π of efficient ϵ-IND-CDP two-party protocols achieving accuracy a for a functionality F, it must be that there is an infinite subfamily $\Pi' \subset \Pi$ such that Π' is an infinite family of $(\epsilon, \delta =$ negligible$)$-DP protocols achieving accuracy a for the functionality F.*

Proof: Suppose that one-way functions do not exist, and yet the statement of the theorem does not hold with respect to some family $\Pi = \{\pi_k\}$ of ϵ-IND-CDP two-party protocols achieving accuracy a for a functionality F. Then there must exist some constant $c > 0$ and a number k_0 such that for all $k > k_0$, it is the case that π_k achieves accuracy $a - k^{-c}$ for F (this follows by assumption) and yet the protocol is not $(\epsilon + k^{-c}, \delta = k^{-c})$-DP.

Without loss of generality, let us assume that the two parties are named Alice and Bob, and it is Alice's privacy that is violated with respect to the $(\epsilon + k^{-c}, \delta = k^{-c})$-DP definition. Let us denote Bob's view in the protocol π on inputs (x, y) by the random variable $\pi(x, y)$, and let us write π to denote π_k to ease our notation. Then it must be that there exist two neighboring inputs x_0, x_1 for Alice, an input y for Bob, and a subset S of views of Bob such that:

$$\Pr[\pi(x_0, y) \in S] > e^{(\epsilon + k^{-c})} \Pr[\pi(x_1, y) \in S] + k^{-c}.$$

Our goal is to show that Alice's privacy is also violated with respect to the ϵ-IND-CDP definition, thereby reaching a contradiction. To show this, we will need to replace the test subset S with an efficiently recognizable test. We will do this in several steps. (Note that we do not attempt to optimize choice of parameters in this proof.) First, let us consider the following subset of S:

$$S' = \left\{ v \in S \;\middle|\; \Pr[\pi(x_0, y) = v] > e^{(\epsilon + k^{-c})} \Pr[\pi(x_1, y) = v] \right\}.$$

It is immediate that we have:

$$\Pr[\pi(x_0, y) \in S'] > e^{(\epsilon + k^{-c})} \Pr[\pi(x_1, y) \in S'] + k^{-c}.$$

since views in S that are not in S' contribute only negatively toward satisfying this condition.

Finally, we now consider the potentially larger set T:

$$T = \left\{ v \;\middle|\; \Pr[\pi(x_0, y) = v] > e^{(\epsilon + k^{-c})} \Pr[\pi(x_1, y) = v] \right\}.$$

Again, since $S' \subset T$, and by the definition of T, it is immediate that:

$$\Pr[\pi(x_0, y) \in T] > e^{(\epsilon + k^{-c})} \Pr[\pi(x_1, y) \in T] + k^{-c}.$$

We will now show there exists an efficient test that very closely approximates testing whether or not a view is from T, thus violating the ϵ-IND-CDP condition. To do this, we first define the following function family:

$$f_k(x_0, x_1, y, b, r_A, r_B) = (x_0, x_1, y, \pi_k(x_b, y; r_A, r_B)).$$

In the function above, x_0, x_1 come from the domain of Alice's inputs to the functionality F, while y comes from the domain of Bob's inputs to the functionality F. Note that these domains are of fixed size, independent of k. Above, b is a single bit, and r_A, r_B are the randomness of Alice and Bob, respectively. Let $t(k)$ denote the total length of the input to f_k. By the efficiency of Π, we have that $t(k)$ is bounded by some fixed polynomial in k.

We first define an efficient test that assumes oracle access to a *perfect* inverter I for the function family f_k. That is, we assume that $I(1^k, y)$ samples uniformly from the preimage space $f_k^{-1}(y)$. We will later show that we can relax the perfectness condition, and implement a good-enough inverter efficiently if one-way functions do not exist.

The test E^I on input v does the following: Let $I_b(x_0, x_1, y, v)$ denote the output of executing $I(x_0, x_1, y, v)$ and returning the inverted value of b. E runs $I_b(x_0, x_1, y, v)$ a total of $\ell = 16 \cdot e^{\epsilon} \cdot k^{2c} \cdot t(k)^2$ times, and collects statistics on how often $b = 0$ and $b = 1$. If the number of times $b = 0$ is at least a factor of $e^{(\epsilon + \frac{1}{2} \cdot k^{-c})}$ larger than the number of times $b = 1$, then E outputs 1.

Note that E is a polynomial-time oracle machine. If I is a perfect inverter, by a Chernoff bound, we have that if v is such that

$$\Pr[\pi(x_0, y) = v] \leq e^{\epsilon + \frac{1}{4} \cdot k^{-c}} \Pr[\pi(x_1, y) = v],$$

then the probability that E^I outputs 1 is less than $2^{-2k+t(k)}$, since we have that $e^{\epsilon + \frac{1}{2} \cdot k^{-c}} - e^{\epsilon + \frac{1}{4} \cdot k^{-c}} > \frac{1}{4} k^{-c}$. On the other hand, by another application of Chernoff's bound, for every $v \in T$, the probability that E^I outputs 0 is less than $2^{-2k+t(k)}$. There are at most $2^{t(k)}$ possible views v, and so by a union bound, we have that except with probability 2^{-k}, E^I does not output 1 on any view v such that $\Pr[\pi(x_0, y) = v] \leq e^{\epsilon + \frac{1}{4} \cdot k^{-c}} \Pr[\pi(x_1, y) = v]$, but E^I outputs 1 on all $v \in T$.

Thus, except with probability 2^{-k}, we have that:

$$\Pr[E^I(\pi(x_0, y)) = 1] > e^{(\epsilon + \frac{1}{4} \cdot k^{-c})} \Pr[E^I(\pi(x_1, y)) = 1] + k^{-c}.$$

And so, overall for large enough k, we have that:

$$\Pr[E^I(\pi(x_0, y)) = 1] > e^{(\epsilon + \frac{1}{4} \cdot k^{-c})} \Pr[E^I(\pi(x_1, y)) = 1] + \frac{1}{2} \cdot k^{-c}.$$

Finally, to make this an efficient attack, we replace the perfect inverter with an efficient inverter guaranteed to exist for infinitely many k if one-way functions do not exist [IL89]. Such an inverter I' is polynomial-time and has the property that for infinitely many k, the following two distributions D_1 and D_2 have statistical distance at most $1/(2d \cdot \ell \cdot k^{2c})$, where d is the size of the domain of Alice squared times the size of the domain of Bob:

$$D_1 = (f_k(\alpha_0, \alpha_1, \beta, b, r_A, r_B), (\alpha_0, \alpha_1, \beta, b, r_A, r_B))$$

and

$$D_2 = (v = \pi_k(\alpha_b, y), I'(\alpha_0, \alpha_1, \beta, v)),$$

where in the above distributions, $\alpha_0, \alpha_1, \beta, b, r_A, r_B$ are all chosen uniformly from their respective domains. Note that the distribution D_1 is identical to the distribution $\hat{D}_2 = (v = \pi_k(\alpha_b, y), I(\alpha_0, \alpha_1, \beta, v))$, where I is the perfect inverter.

By a union bound over choice of $\alpha_0, \alpha_1, \beta$, we immediately obtain that the following two distributions D'_1, D'_2 have statistical distance at most $1/(2 \cdot \ell \cdot k^{2c})$:

$$D'_1 = (f_k(x_0, x_1, y, b, r_A, r_B), b)$$

and

$$D'_2 = (v = \pi_k(x_b, y), I'_b(x_0, x_1, y, v)),$$

where in the above distributions only b, r_A, r_B are chosen uniformly from their respective domains.

Thus, since $E^{I'}$ calls I'_b exactly ℓ times, by a union bound we have that for infinitely many k,

$$\Pr[E^{I'}(\pi(x_0, y)) = 1] > e^{(\epsilon + \frac{1}{4} \cdot k^{-c})} \Pr[E^{I'}(\pi(x_1, y)) = 1] + \frac{1}{4} \cdot k^{-c}.$$

This contradicts the ϵ-IND-CDP property of Π, finishing the proof. ∎

References

[BN10] Hai Brenner and Kobbi Nissim. Impossibility of differentially private universally optimal mechanisms. In Trevisan [Tre10], pages 71–80.

[BNO08] Amos Beimel, Kobbi Nissim, and Eran Omri. Distributed private data analysis: Simultaneously solving how and what. In Wagner [Wag08], pages 451–468.

[CK91] Benny Chor and Eyal Kushilevitz. A zero-one law for Boolean privacy. *SIAM J. Discrete Math.*, 4(1):36–47, 1991.

[CKL03] Ran Canetti, Eyal Kushilevitz, and Yehuda Lindell. On the limitations of universally composable two-party computation without set-up assumptions. In *EUROCRYPT*, pages 68–86, 2003.

[DKM$^+$06] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: privacy via distributed noise generation. In *Advances in Cryptology—EUROCRYPT 2006*, volume 4004 of *Lecture Notes in Computer Science*, pages 486–503. Springer, 2006.

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Shai Halevi and Tal Rabin, editors, *Theory of Cryptography, Third Theory of Cryptography Conference, TCC 2006*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284. Springer, 2006.

[DMT07] Cynthia Dwork, Frank McSherry, and Kunal Talwar. The price of privacy and the limits of LP decoding. In David S. Johnson and Uriel Feige, editors, *STOC*, pages 85–94. ACM, 2007.

[DN03] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *PODS*, pages 202–210. ACM, 2003.

[DN04] Cynthia Dwork and Kobbi Nissim. Privacy-preserving datamining on vertically partitioned databases. In Matthew K. Franklin, editor, *Advances in Cryptology—CRYPTO 2004*, volume 3152 of *Lecture Notes in Computer Science*, pages 528–544. Springer, 2004.

[DNR$^+$09] Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil P. Vadhan. On the complexity of differentially private data release: efficient algorithms and hardness results. In Mitzenmacher [Mit09], pages 381–390.

[Dwo06] Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.

[Dwo11] Cynthia Dwork. A firm foundation for private data analysis. *Commun. ACM*, 54(1):86–95, 2011.

[DY08] Cynthia Dwork and Sergey Yekhanin. New efficient attacks on statistical disclosure control mechanisms. In Wagner [Wag08], pages 469–480.

[GMW87] Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play any mental game or a completeness theorem for protocols with honest majority. In Alfred V. Aho, editor, *STOC*, pages 218–229. ACM, 1987.

[GRS09] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In Mitzenmacher [Mit09], pages 351–360.

[GS10] Mangesh Gupte and Mukund Sundararajan. Universally optimal privacy mechanisms for minimax agents. In Jan Paredaens and Dirk Van Gucht, editors, *PODS*, pages 135–146. ACM, 2010.

[HNRR06] Danny Harnik, Moni Naor, Omer Reingold, and Alon Rosen. Completeness in two-party secure computation: A computational view. *J. Cryptology*, 19(4):521–552, 2006.

[HT10] Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In Schulman [Sch10], pages 705–714.

[IL89] Russell Impagliazzo and Michael Luby. One-way functions are essential for complexity based cryptography (extended abstract). In *FOCS*, pages 230–235. IEEE, 1989.

[IPS08] Yuval Ishai, Manoj Prabhakaran, and Amit Sahai. Founding cryptography on oblivious transfer—efficiently. In Wagner [Wag08], pages 572–591.

[Kil88] Joe Kilian. Founding cryptography on oblivious transfer. In Janos Simon, editor, *STOC*, pages 20–31. ACM, 1988.

[Kil00] Joe Kilian. More general completeness theorems for secure two-party computation. In *STOC*, pages 316–324, 2000.

[KRSU10] Shiva Prasad Kasiviswanathan, Mark Rudelson, Adam Smith, and Jonathan Ullman. The price of privately releasing contingency tables and the spectra of random matrices with correlated rows. In Schulman [Sch10], pages 775–784.

[Mit09] Michael Mitzenmacher, editor. *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*. ACM, 2009.

[MMP$^+$10] Andrew McGregor, Ilya Mironov, Toniann Pitassi, Omer Reingold, Kunal Talwar, and Salil P. Vadhan. The limits of two-party differential privacy. In Trevisan [Tre10], pages 81–90.

[MPRV09] Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil P. Vadhan. Computational differential privacy. In Shai Halevi, editor, *CRYPTO*, volume 5677 of *Lecture Notes in Computer Science*, pages 126–142. Springer, 2009.

[Sch10] Leonard J. Schulman, editor. *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*. ACM, 2010.

[Tre10] Luca Trevisan, editor. *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*. IEEE Computer Society, 2010.

[UV11] Jonathan Ullman and Salil P. Vadhan. PCPs and the hardness of generating private synthetic data. In Yuval Ishai, editor, *TCC*, volume 6597 of *Lecture Notes in Computer Science*, pages 400–416. Springer, 2011.

[Wag08] David Wagner, editor. *Advances in Cryptology—CRYPTO 2008, 28th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2008. Proceedings*, volume 5157 of *Lecture Notes in Computer Science*. Springer, 2008.

[Yao82] Andrew Chi-Chih Yao. Protocols for secure computations (extended abstract). In *23rd Annual Symposium on Foundations of Computer Science (FOCS 1982)*, pages 160–164. IEEE, 1982.